



HAL
open science

Communication Améliorée dans un monde virtuel

Eric Varadaradjou

► **To cite this version:**

Eric Varadaradjou. Communication Améliorée dans un monde virtuel. domain_other. Télécom ParisTech, 2007. English. NNT: . pastel-00004077

HAL Id: pastel-00004077

<https://pastel.hal.science/pastel-00004077>

Submitted on 9 Jan 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



ENST



EDITE

version du 8 juin 2008

Communication améliorée dans un monde virtuel

Éric Saravanane VARADARADJOU

Directeurs de thèse :

**Philippe DAX
Alain GRUMBACH**

Ecole Nationale Supérieure des Télécommunications

Remerciements

Les remerciements... Une partie encore plus difficile à rédiger que le reste... où le doctorant doit entreprendre de résumer en quelques lignes toute l'aide qu'on lui a apporté durant ses années de thésard.

Je tiens tout d'abord à remercier vivement mes directeurs de thèse Monsieur Alain GRUMBACH, professeur à l'ENST, et Monsieur Philippe DAX, Directeur d'études au département Informatique et Réseaux, qui ont su m'aider à démêler l'écheveau des différentes problématiques. Souvent désorienté, parfois perdu, ils ont su me recadrer tout en encourageant mes initiatives et en valorisant mes idées. Sans oublier, bien sûr, la confiance qu'ils m'ont accordée ainsi que l'immense patience dont ils ont fait preuve tout au long de mon travail de thèse.

Je remercie vivement Monsieur Pierre NUGUES, Professeur à l'Université de Lund (Suède), dont les travaux, l'environnement Ulysse en particulier, ont très largement inspiré cette recherche. Une majeure partie de ce travail peut être considérée comme un prolongement d'investigations amorcées par Pierre NUGUES. Je remercie Monsieur Christophe CHAILLOU, Professeur à POLYTECH'LILLE, pour ses critiques très constructives relatives à la conjonction : communication - virtuel. Je remercie Monsieur Jean-Claude MARTIN, chercheur au LIMSI, dont les travaux éclairants relatifs aux modalités m'ont permis d'aborder et de fusionner différents axes de recherche.

Je remercie les différentes personnes qui ont participé de près ou de loin à l'écriture de ce manuscrit : Annie Danzart et Christine Potier pour les différentes lectures et surtout pour leurs remarques nombreuses et constructives, Ayline pour ses conseils et surtout Alain pour ses trèèèèèès nombreuses relectures.

Pour reprendre un ordre plus chronologique, je voudrais remercier deux camarades devenus amis et qui ont rendu mon séjour en C255 des plus agréables, Lucile, la lanceuse de peluches (ange ou démon ?) ainsi que Thomas (ou comment communiquer avec les tables). Ils m'ont permis à de nombreuses reprises de prendre du recul sur mes recherches afin de mieux les comprendre.

Je remercie Céline pour son soutien, sa confiance et sa très grande gentillesse. Enfin merci à Hoa qui m'a soutenu durant la période de solitude que représente l'écriture du manuscrit.

Je souhaite ensuite remercier l'équipe des mathématiciens «fous» : Olivier, Irène et Antoine; fous par leur génie à comprendre et résoudre ces problèmes incompréhensibles au commun des mortels et fous pour leur passion immodérée du chocolat. Je les remercie pour toutes ces invitations aux petits goûters qui m'ont fourni maintes occasions de faire des pauses bien nécessaires. Merci notamment à Serge, à Eric et aux membres de l'ancien groupe I3 pour nos discussions et nos débats. Je souhaite aussi remercier de manière générale toutes les personnes du département Informatique et Réseaux ainsi que les nombreux élèves de l'ENST dont Julien (Juju) qui m'a enseigné l'art d'enseigner.

De manière plus personnelle, je tiens à remercier Adeline pour nos nombreuses discussions, Bennyben et Johanne pour leur éternelle bonne humeur, Nanou et Vijay pour leurs coups de main en anglais et enfin Ada (puffetta) pour ses coups de pieds.

Enfin, je n'oublierai jamais les amis perdus des jardins de Fleming et les membres de l'association Flemnet, dont la compagnie d'Idriss (Iorikyu) et Neerwan (Sly) pour ses soirées VIP.

Résumé

Cette recherche s'inscrit à la croisée des chemins de la réalité virtuelle et de l'intelligence artificielle. Elle concerne la communication. Considérons deux humains évoluant dans un même monde virtuel et communiquant à ce sujet. Deux questions se posent :

- *Quelles sont les caractéristiques spécifiques de la **communication** dans un **monde virtuel** ?*
- *Quelles **améliorations** de la communication sont réalisables grâce au **virtuel** ?*

Les améliorations envisagées portent sur la communication à travers :

- les interlocuteurs : deux humains connectés à un monde virtuel, ou un humain communiquant avec son avatar ;
- le contenu des actes de communication ;
- les modalités utilisées.

Cette recherche se situe dans la droite ligne des travaux précurseurs de P. Nuges qui a réalisé le système de navigation *Ulysse* permettant à un humain de communiquer dans un monde virtuel avec son avatar par le biais d'un agent conversationnel. Notre approche poursuit cette démarche en étendant :

- le type de communication : entre deux humains ;
- les modalités utilisées : complémentarité entre la modalité textuelle et la modalité déictique.

Concernant la communication entre humains, nous étudions les apports d'un agent conversationnel qui assiste le dialogue entre les interlocuteurs en y apportant ses compétences en termes d'informations relatives au monde virtuel. Ces apports permettent par exemple de :

- détecter, montrer puis lever des ambiguïtés éventuelles apparaissant dans un acte de communication, par l'apport d'informations contextuelles ;
- assister, augmenter, automatiser la réalisation d'une requête.

Le manuscrit débute par une synthèse des différents modèles de communication. Il se poursuit par une analyse de modes de communication présents

dans les mondes virtuels. Puis nous proposons un ensemble de fonctionnalités qui peuvent enrichir la communication en utilisant les possibilités apportées par le virtuel. Enfin nous décrivons et discutons de l'implémentation de ces fonctionnalités ainsi que celles de l'agent conversationnel au sein de l'environnement VREng développé à l'ENST sous la direction de P. Dax.

Abstract

This research is at crossroads of Virtual Reality and Artificial Intelligence and focuses on communication challenges. Let's consider two humans evolving in the same virtual world and communicating on it; we may ask ourselves 2 questions :

- *what are the specific features of **communication** in a **virtual world** ?*
- *what kind of **communication improvement** can be expected from **virtual reality** ?*

In particular, improvement focuses on :

- interlocutors : two people connected in a virtual world, or a human communicating with his avatar ;
- content of the act of communicating ;
- communication modes.

Starting point of our research is **P. Nugues'** precursor work, and his navigation system named **Ulysse** that allowing human to communicate in a virtual world through his avatar by means of a conversational agent. Our research considers the following elements :

- type of communication : between two people ;
- modes used to communicate : integration of textual and deictic modes.

More in depth, we identify conversational agent benefits to dialogue between humans. The agent integrates information in the virtual world, and consequently enables enhancements to communication such as :

- detection, display, and removal of possible ambiguities in communication by providing contextual information ;
- assistance and support to automatic query execution.

Overview of the research : the manuscript begins with a summary of communication models, and then deals with the analysis of communication patterns in the virtual worlds and the proposal on a set of features to enhance communication. Finally, we explain the implementation of a conversational agent and of communication-enhancing features in the VREng environment, that was developed at the ENST under the direction of **P. Dax**.

Riassunto

La ricerca si colloca nell'intersezione tra la realtà virtuale e l'intelligenza artificiale e riguarda la comunicazione. Infatti, considerando due esseri umani che si evolvono in un medesimo mondo virtuale e comunicano a questo proposito, ci si pongono due domande :

- *Quali sono le caratteristiche specifiche della **comunicazione** in un **mondo virtuale**?*
- *Quali **miglioramenti** della comunicazione si possono realizzare grazie alla realtà **virtuale**?*

I miglioramenti ipotizzati si focalizzano sulla comunicazione, tramite :

- gli interlocutori : due esseri umani collegati in un mondo virtuale, o un singolo che comunica col suo avatar ;
- il contenuto degli atti di comunicazione ;
- le modalità di comunicazione.

Questa ricerca prosegue nella direzione del lavoro precursore di **P. Nugues**, ideatore del sistema di navigazione **Ulysse**, che permette all'essere umano di comunicare in un mondo virtuale col sua avatar tramite un agente conversazionale. La ricerca amplia questo approccio considerando :

- il tipo di comunicazione : tra due esseri umani ;
- le modalità utilizzate : complementarità tra la modalità testuale e quella deittica.

Riguardo alla comunicazione tra esseri umani, viene studiato il contributo che un agente conversazionale dà al dialogo tra gli interlocutori, fornendo informazioni sul mondo virtuale. Grazie a questo contributo, è possibile ad esempio :

- scoprire, mostrare e quindi eliminare le eventuali ambiguità in un atto di comunicazione, grazie all'apporto di informazioni contestuali ;
- assistere, incrementare ed automatizzare la realizzazione di una richiesta.

Il manoscritto si apre con una sintesi dei vari modelli di comunicazione. In seguito vengono analizzate le modalità di comunicazione nei mondi virtuali

e proposte delle funzionalità che possono arricchirla. Infine, viene descritta l'implementazione di queste funzionalità e di quelle dell'agente conversazionale all'interno dell'ambiente virtuale VREng, sviluppato all'ENST sotto la direzione di P. Dax.

Table des matières

1	Introduction	19
2	La communication	23
2.1	Introduction	24
2.2	Évolution de la communication	25
2.3	Etude de la communication	27
2.4	Les modèles de communication	29
2.4.1	Le modèle de Shannon et Weaver	29
2.4.2	Le modèle de Lasswell	32
2.4.3	Le modèle de Gerbner	34
2.4.4	Le modèle de Newcomb	37
2.4.5	Le modèle de Westley et Mac Lean	38
2.4.6	Le modèle de Riley et Riley	39
2.5	Conclusion	41
2.5.1	Les problèmes de communication	41
2.5.2	Le récepteur	42
2.5.3	La forme agit sur le contenu	42
2.5.4	Les lois de la bonne communication	42
2.5.5	Conclusion	43
3	La réalité virtuelle	45
3.1	Introduction	46
3.1.1	Dispositifs immersifs	47
3.1.2	Dispositifs semi-immersifs	47
3.1.3	Dispositifs non-immersifs	48
3.2	Applications	49
3.3	Apports et limites des mondes virtuels	49
3.3.1	Sequence World : visualisation d'indices	50
3.3.2	Newton World : l'influence de la gravité	51
3.3.3	Limitations des mondes virtuels	51
3.4	Étude des communications dans les mondes virtuels	53
3.4.1	Introduction de l'étude	55
3.4.2	Les Environnements Virtuels	55

3.4.3	Dive : Distributed Interactive Virtual Environment . . .	55
3.4.4	3DWebmaster : le langage SCL	58
3.4.5	Ulysse : agent conversationnel interactif	60
3.4.6	VREng : Virtual Reality Engine	63
3.4.7	Diamond Park : système social virtuel	64
3.4.8	Active World : l'un des premiers EV	67
3.4.9	Open Croquet : mondes virtuels collaboratifs	69
3.4.10	Habbo hotel : un monde en ligne	71
3.4.11	Spin-3D : pour un travail collaboratif	73
3.4.12	LEA : LIMSI Embodied Agent	75
3.4.13	Divers mondes virtuels	77
3.4.14	Synthèse de l'étude	81
3.4.15	La RV et la communication	84
3.5	Conclusion	88
4	La communication améliorée	89
4.1	Le concept d'amélioration	90
4.1.1	Définitions	90
4.1.2	Nos champs d'actions	91
4.1.3	Exemples d'améliorations	96
4.1.4	Conclusion	99
4.2	Mise en œuvre des améliorations	100
4.2.1	L'agent	100
4.2.2	Cas d'utilisation	104
4.2.3	Les modes opératoires de l'agent	111
4.2.4	Réalisation des améliorations	114
4.2.5	Conclusion	119
4.3	Conclusion	120
5	Les implémentations	121
5.1	VREng : «Virtual Reality Engine»	122
5.1.1	Structure de VREng	123
5.1.2	Structure des données	124
5.1.3	Mise en place de l'agent	127
5.1.4	Conclusion	131
5.2	Implémentation des méthodes	132
5.3	Etayer les informations	132
5.3.1	Améliorations textuelles (T)	132
5.3.2	Améliorations déictiques (D)	142
5.3.3	Améliorations visuelles (V)	147
5.3.4	Conclusion sur les améliorations	151
5.4	Prévenir les ambiguïtés	151
5.4.1	Prévention des ambiguïtés géographiques (P1)	151
5.4.2	Prévention des ambiguïtés visuelles (P2)	152

5.4.3	Conclusion sur la prévention des ambiguïtés	153
5.5	Répondre aux attentes	154
5.5.1	Réponse aux ordres (R1)	154
5.5.2	Réponse aux questions (R2)	155
5.6	Conclusion	156
6	Evaluation	159
6.1	Principes de Newcomb	159
6.2	Scénarios sur la communication améliorée	160
6.2.1	Scénario sans ACI : S1	160
6.2.2	Scénario avec ACI : S2	161
6.2.3	Conclusion	163
6.3	Proposition pour l'évaluation de l'ACI	164
6.4	Prototype d'ACI	165
7	Conclusions et perspectives	169
8	Annexes	173
8.1	Autorisations Images	173
8.2	Analyse du mode de «picking»	177
8.3	Positionnement automatique	180
8.4	Transmission d'images à travers un réseau	182
	Bibliographie	184
	Index	201

Table des figures

1.1	<i>Un acte de communication</i>	20
1.2	<i>Une possibilité du support virtuel</i>	20
1.3	<i>Une autre possibilité du support virtuel</i>	21
2.1	<i>Illustration de la communication pyramidale</i>	25
2.2	<i>La communication de vis à vis</i>	25
2.3	<i>La communication téléphonique</i>	26
2.4	<i>La communication par visioconférence</i>	26
2.5	<i>La communication améliorée</i>	27
2.6	<i>Modèle communicationnel de Shannon et Weaver</i>	30
2.7	<i>Modèle communicationnel de Shannon-Weaver adapté à la RV</i>	31
2.8	<i>Modèle communicationnel de Lasswell</i>	32
2.9	<i>Études associées aux composantes de la communication</i>	33
2.10	<i>Modèle communicationnel de Gerbner</i>	35
2.11	<i>Modèle communicationnel de Newcomb</i>	37
2.12	<i>Modèle communicationnel de Westley et Mac Lean</i>	38
2.13	<i>Modèle communicationnel sur l'émission d'un message</i>	39
2.14	<i>Modèle communicationnel de Riley et Riley</i>	40
3.1	<i>Exemples de mondes virtuel</i>	47
3.2	<i>Dispositif immersif : un CAVE</i>	47
3.3	<i>Dispositif semi-immersif : un Workbench</i>	48
3.4	<i>Dispositif non-immersif : un phantom</i>	49
3.5	<i>Un arbre de cubes</i>	50
3.6	<i>Vue sur un monde de SequenceWorld</i>	51
3.7	<i>Vue sur un monde de Newton World</i>	52
3.8	<i>Illustrations d'une limitation des MV</i>	52
3.9	<i>Vues du monde de DIVE</i>	56
3.10	<i>Image d'une fenêtre de XChat</i>	57
3.11	<i>Vues de SuperScape-3D Webmaster</i>	59
3.12	<i>Vue sur un monde d'Ulysse</i>	61
3.13	<i>Exemple d'une requête</i>	62
3.14	<i>Vue du monde de VREng</i>	63

3.15	<i>Vues sur un monde de Diamond Park</i>	65
3.16	<i>Vue sur un monde d'Active World</i>	67
3.17	<i>Vue sur un monde d'Open Croquet</i>	69
3.18	<i>Vue sur un monde de Habbo</i>	71
3.19	<i>Vue sur Argonaute 3D</i>	74
3.20	<i>Vue sur le système DaftLEA</i>	76
3.21	<i>Vue sur un monde de Second Life</i>	78
3.22	<i>Vue sur un monde des sims online</i>	79
3.23	<i>Vue sur un monde de Village 3D</i>	80
3.24	<i>Vue sur un monde de VR4all</i>	81
3.25	<i>Divers MMOG</i>	82
4.1	<i>Histogramme sur la longueur des phrases</i>	93
4.2	<i>Schéma fonctionnel d'une modalité</i>	105
4.3	<i>La modalité textuelle</i>	106
4.4	<i>La modalité textuelle améliorée</i>	106
4.5	<i>La modalité déictique</i>	107
4.6	<i>La modalité déictique améliorée</i>	107
4.7	<i>Incompréhension par l'ACI</i>	112
4.8	<i>Amélioration par l'ACI</i>	112
4.9	<i>Levée d'ambiguïté par l'ACI</i>	113
4.10	<i>Réponse de l'ACI</i>	113
4.11	<i>Action de l'ACI</i>	114
4.12	<i>Exemple d'amélioration du message transmis</i>	116
4.13	<i>Utilisation de la désignation améliorée</i>	117
5.1	<i>Un musée virtuel sous VREng</i>	122
5.2	<i>Schéma général de VREng</i>	123
5.3	<i>Schéma de VREng avant l'agent</i>	127
5.4	<i>Schéma de VREng avec l'agent</i>	128
5.5	<i>Exemple de scènes</i>	135
5.6	<i>Illustration de la perception relative au positionnement</i>	138
5.7	<i>Désignation d'un ensemble de livres</i>	144
5.8	<i>Visualisation du point de vue de la caméra</i>	146
5.9	<i>Visualisation d'un point de vue</i>	148
5.10	<i>Visualisation du point de vue de la caméra</i>	150
6.1	<i>Utilisateur : «Ouvre !»</i>	166
6.2	<i>Utilisateur : «Ferme cette porte !»</i>	166
6.3	<i>Utilisateur : «Ouvre cette porte !»</i>	167
8.1	<i>Le champ de vision sous OpenGL</i>	180
8.2	<i>La recherche de positionnement</i>	181
8.3	<i>Les tampons d'affichage sous OpenGL</i>	183

Liste des tableaux

4.1	Tableau des statistiques	92
5.1	Résultat des incréments successives	136
5.2	Classement par pertinence	137
8.1	Les images sous licence GPL	174
8.2	Les images sous Licence de diffusion libre	175
8.3	Les images avec autorisations spécifiques	176

1 Introduction

Dans l'ère actuelle de l'information, les systèmes et les supports de communication se diversifient et se généralisent. Des outils de communication innovants apparaissent. La réalité virtuelle est un contexte favorable au développement de tels outils. Afin d'illustrer notre problématique, considérons un scénario où deux visiteurs d'un musée virtuel engagent une conversation.

Exemple de scénario

- Cadre : visite d'une exposition virtuelle d'œuvres d'art ;
- Visiteurs : A et E ;
- Support de communication : boîte de dialogue (*chat*) et outil de désignation au sein d'un monde virtuel.

A a terminé la visite. E est encore à l'exposition.

A souhaite retrouver E dans un cadre convivial.

A se dirige vers un café pour attendre E.

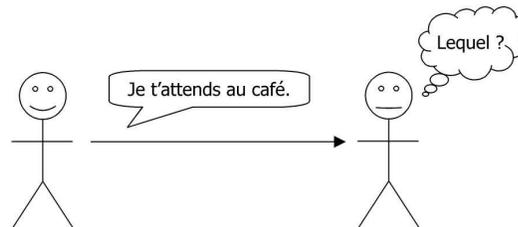
Une conversation s'engage entre A et E :

01. A -> E : J'ai terminé la visite; je t'attends au café.
02. E -> A : D'accord, celui qui est à côté de la fontaine ?
03. A -> E : Non, celui qui est près de la bibliothèque.
04. E -> A : Je ne vois pas.
05. A -> E : Attends, je vais voir son nom.
{déplacement de A vers le café}
Il s'appelle: Chez-Paul ! Tu le connais ?
06. E -> A : Non, ce nom ne me dit rien.
07. A -> E : Tu préfères aller dans un autre café ?
08. E -> A : Oui, je préfère Chez-Pierre.
09. A -> E : Où est le café Chez-Pierre ?
10. E -> A : Il est près de la fontaine,
tu ne peux pas le rater.
11. A -> E : D'accord, j'y vais.

Légende :

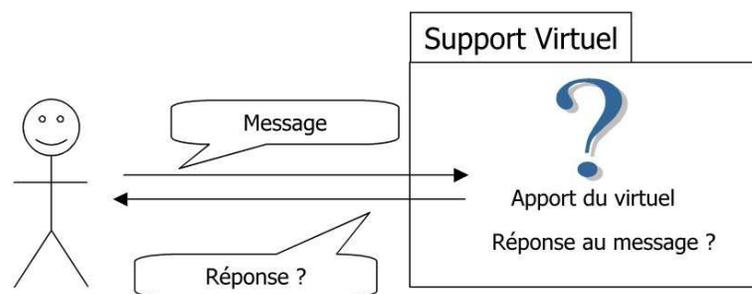
- «X -> Y : M» signifie X envoie le message M à Y ;
- «{Action}» signifie Action commandée par l'utilisateur et effectuée par l'avatar de l'utilisateur.

Nous pouvons noter dans cet exemple, de nombreuses imperfections du langage comme à la ligne 01 où A, qui souhaite retrouver E, n'a pas connaissance des différents cafés. Ceci entraîne une ambiguïté sur le lieu de rendez-vous. Or, dans le contexte d'une évolution constante des technologies de l'information et de la communication, nous aurions pu penser à une levée de certaines ambiguïtés, avec pour conséquence plus de clarté dans les communications. Nous voyons dans la figure 1.1 que les actes de communication sont transmis sans traitement. C'est aux interlocuteurs de prévoir ces ambiguïtés. Les recherches ont été nombreuses mais ces études se sont en majeure partie intéressées aux Interactions Homme-Machine. Elles n'abordent pas les **communications interpersonnelles médiatisées par la machine**.

FIG. 1.1 – *Un acte de communication*

De ce fait, la problématique abordée peut être ponctuée par la question :

Quelles sont les possibilités nouvelles, spécifiques, permettant d'enrichir la communication entre locuteurs, conséquentes du fait que le monde considéré est virtuel ? (cf. figure 1.3)

FIG. 1.2 – *Une possibilité du support virtuel*

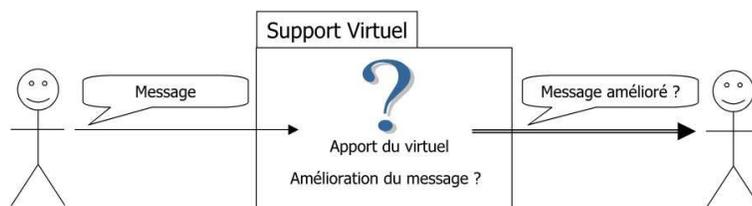


FIG. 1.3 – Une autre possibilité du support virtuel

Motivations

Au carrefour de deux grands domaines que sont la **réalité virtuelle** et la **communication**, nous avons tenté d'extraire des aspects innovants relatifs au contexte de la communication entre locuteurs. Par exemple, la communication revêt un caractère important pour de nombreux travaux collaboratifs. Elle est à la source des possibilités qui permettent aux interlocuteurs de se comprendre et d'atteindre des objectifs communs. Mais le constat a été de remarquer que ce type de communication restait peu étudié.

L'objectif du projet ACI (Agent Conversationnel Intelligent) pour la plate-forme VREng (Virtual Reality Engine) est d'intervenir dans les actes de communication et de fournir aux utilisateurs des moyens d'améliorer la communication suivant trois axes :

- **étayer**, préciser les informations contenues dans les messages en analysant les contextes liés au message.
- **prévenir**, anticiper les mauvaises compréhensions du message par une analyse préventive du contexte virtuel.
- **répondre** à certaines attentes de l'utilisateur.

Ainsi nous verrons comment *via* notre agent, la conversation entre les deux visiteurs de l'exposition virtuelle peut être améliorée.

Contributions

Les contributions de ce mémoire se situent sur plusieurs plans.

La liaison entre le domaine de la communication et le domaine de la réalité virtuelle est peu étudiée. Les différentes questions soulevées par la communication trouvent parfois des réponses dans le domaine de la réalité virtuelle (cf. Chapitre 3, Section 3.4.15, page 84). En nous inspirant d'approches en réalité virtuelle, nous avons pu proposer diverses formes d'améliorations à la communication (cf. Chapitre 4, Section 4.1.3, page 96). A l'aide d'analyses langagières (cf. Chapitre 4, Section 4.1.2, page 91) nous

avons pu mettre au point un système de traitements capable d'intervenir au niveau des messages transmis entre les utilisateurs afin de les améliorer (cf. Chapitre 4, Section 4.2.3, page 111). De notre point de vue, une des innovations est l'inscription des problèmes spécifiques à la communication dans les mondes virtuels suivant un contexte élargi. En d'autres termes, certaines problématiques de la communication générale sont modifiées par le fait que les supports ont évolué.

La mise en place d'un **agent conversationnel intelligent** au sein de l'environnement VREng a permis de réaliser des méthodes d'amélioration de la communication. Par exemple, au moyen d'une analyse langagière et d'un processus d'inférence, nous avons doté l'agent de comportements qui sont fonctions des informations contenues dans les messages. La portée de ses actions est interne (modification du monde) et externe (interaction avec l'utilisateur).

Organisation du mémoire

Le présent manuscrit s'articule autour de trois grandes parties qui répondent avant tout à une volonté de cohérence :

- Partie 1 : En premier lieu, nous nous sommes intéressés à l'étude de la **communication** en général afin de comprendre les concepts, les modèles et les questions spécifiques à ce domaine (cf. Chapitre 2). Sur cette base, nous avons pu étudier les apports possibles au sein de **mondes virtuels**. Par la suite, l'analyse des communications dans les mondes virtuels nous a permis de comprendre les possibilités permettant d'obtenir une amélioration des actes de communications (cf. Chapitre 3).
- Partie 2 : La deuxième partie énonce les bases du concept de **communication améliorée** (cf. Chapitre 4) et présente les idées apparues dans les analyses précédentes de manière plus concrète. Par la suite, nous développons les possibilités de l'**agent conversationnel intelligent** ainsi que son champ d'action. Nous exposons à travers des exemples, les différentes possibilités d'améliorations qui se classent en trois catégories : **étayer**, **prévenir** et **répondre**.
- Partie 3 : Enfin, nous détaillons les différentes **implémentations des techniques** que nous avons employées pour élaborer nos améliorations (cf. Chapitre 5). L'étude est illustré par un couple de scénarios. Elle se termine par une proposition pour la validation et une évaluation de l'ACI (cf. Chapitre 6).

2 La communication

Sommaire

2.1	Introduction	24
2.2	Évolution de la communication	25
2.3	Etude de la communication	27
2.4	Les modèles de communication	29
2.4.1	Le modèle de Shannon et Weaver	29
2.4.2	Le modèle de Lasswell	32
2.4.3	Le modèle de Gerbner	34
2.4.4	Le modèle de Newcomb	37
2.4.5	Le modèle de Westley et Mac Lean	38
2.4.6	Le modèle de Riley et Riley	39
2.5	Conclusion	41
2.5.1	Les problèmes de communication	41
2.5.2	Le récepteur	42
2.5.3	La forme agit sur le contenu	42
2.5.4	Les lois de la bonne communication	42
2.5.5	Conclusion	43

La communication est plus rare que le bonheur, plus fragile que la beauté.
[Emmanuel Mounier]

Dans cette partie de notre état de l'art, nous allons étudier le domaine pluridisciplinaire de la communication afin d'approfondir notre connaissance sur le processus et la nature de la communication et mettre en exergue les principales questions de notre problématique. Nous présenterons après une introduction générale sur la communication, notre propre vision de l'évolution de la communication. Par la suite, nous étudierons à l'aide de différents modèles communicationnels connus les problèmes soulevés. Enfin, nous résumerons l'ensemble des questions et problèmes présents.

2.1 Introduction

Depuis l'antiquité, la communication entre les êtres représente un sujet d'étude passionnant pour les philosophes. Néanmoins, c'est relativement récemment qu'elle est devenue une matière d'investigation scientifique. De nos jours, la communication est reconnue comme une discipline établie, rigoureuse avec des départements universitaires, des programmes doctoraux, des revues et organisations scientifiques.

La communication est à la base de l'interaction et de l'organisation sociale. Sans communication, nous sommes isolés¹. La communication crée la coopération, elle permet la relation et l'enrichissement de nos connaissances. Elle se présente sous différentes formes. La forme orale est la plus connue. Ses différentes formes ont pour but commun de faire passer des informations, des messages à l'aide d'un langage commun.

Lors de la «communication», les informations vont circuler entre les communicants. Ces communicants apprennent ainsi de nouvelles informations, non seulement contenues dans l'information elle-même, mais aussi dans le contexte de l'information. L'objectif de la communication peut être assimilé à une mise en commun des informations de chacun.

Dans un monde qui se veut être celui de la communication, mise en commun de nos idées reste pour difficile. Nous sommes loin d'être performants[90]. Pourtant, la diversité des approches et les ouvrages sur le sujet ne manquent pas. L'information, au centre de la communication, reste un sujet complexe et difficile.

Définitions

Pour définir la communication, nous proposerons la définition suivante d'après Chaffee et Berger (1986) : *La science de la communication cherche à comprendre la production, le traitement et les effets des symboles et des systèmes de signes par des théories analysables, contenant des généralisations légitimes permettant d'expliquer les phénomènes associés à la production, aux traitements et aux effets.*

Cette définition permet en outre de se débarrasser des différents contextes communicationnels. En effet, la science de la communication englobe un champ très étendu, que l'on peut diviser en plusieurs niveaux. McQuail [66] a fait une proposition pyramidale de ces différents contextes que nous présentons figure 2.1 :

¹Les bébés ont besoin de communication pour survivre. Le lait et le sommeil ne suffisent pas. La communication est aussi un élément indispensable à la vie. [Bernard Werber]

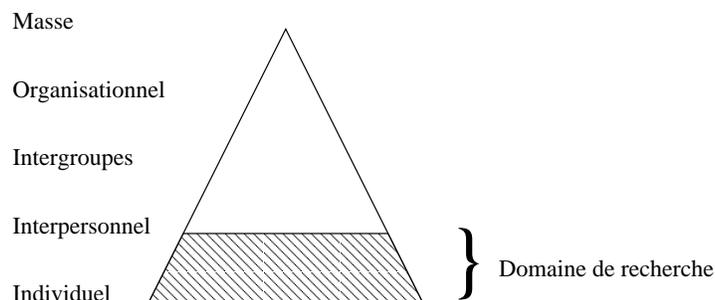


FIG. 2.1 – *Illustration de la communication pyramidale*

Cette représentation pyramidale permet une vision des différents niveaux de communication. Chaque niveau englobe ceux qui se situent en dessous de lui. Ainsi, la communication de masse est le niveau qui se situe au sommet de cette pyramide, et de ce fait concerne tous les autres niveaux. Elle se rapporte à plus de personnes dans le temps et l'espace que les autres, et traite tous les niveaux inférieurs. D'où une certaine prédominance de la recherche dans la communication de masse. Notre recherche quant à elle, traite des communications au niveau interpersonnel qui englobent le traitement du niveau individuel.

2.2 Évolution de la communication

La communication a suivi une évolution parallèle à la technologie. Nombreuses ont été les technologies pour améliorer et accélérer les communications. A ses débuts, l'Homme n'avait pour seul support que la voix, les gestes et ses cinq sens pour parvenir à communiquer et à interagir. Mais ces supports avaient des contraintes importantes. La voix ou les gestes ne pouvaient pas être porteurs d'informations au-delà d'une certaine distance. Les informations ne pouvaient être transmises que de proche en proche (cf. figure 2.2).

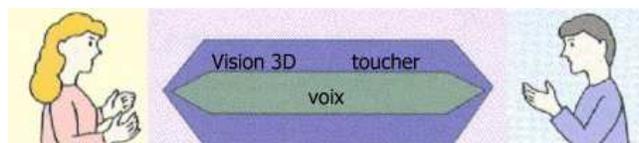


FIG. 2.2 – *La communication de vis à vis*

Les progrès technologiques et plus particulièrement le téléphone ont permis de repousser la barrière de la distance. Nous pouvons dialoguer par la

voix simplement et rapidement avec une personne très éloignée (cf. figure 2.3). Mais cette technique reste un outil de transmission limité². Il ne dispose pas de moyen pour transmettre des informations multimodales telles que les gestes ou les expressions faciales.

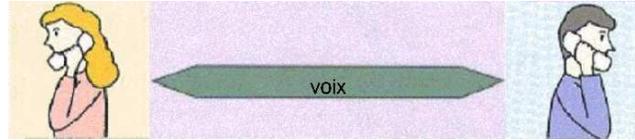


FIG. 2.3 – *La communication téléphonique*

Au cours du temps, les nouvelles technologies de la toile : Internet en premier, nous permettent à l'heure actuelle de diffuser des informations et de les partager très rapidement avec une ou plusieurs personnes à la fois. La nature même des informations transmises s'est amplement diversifiée. Actuellement, la communication par visioconférence s'est largement diffusée notamment à l'aide des messageries instantanées. N'importe quel utilisateur connecté au réseau Internet peut dialoguer et observer à l'aide de caméras un autre utilisateur (cf. figure 2.4).

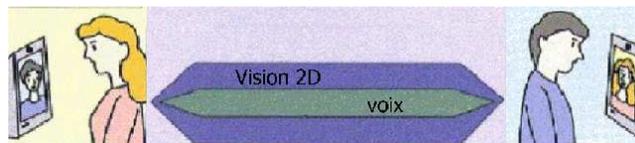


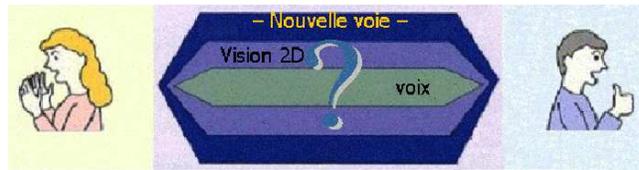
FIG. 2.4 – *La communication par visioconférence*

Ces améliorations technologiques nous permettent de propager des sons, des images, des impressions (avec des gestes, des expressions faciales...). Il est possible d'effectuer des actions collaboratives en temps réel laissant entrevoir de nombreuses possibilités tout aussi bien ludiques que professionnelles (cf. travaux de McQuail [66]), tel les diagnostics collaboratifs. Toutes ces améliorations restent néanmoins limitées.

Que peut-on améliorer dans la communication ? Quelles sont les possibilités nouvelles, permettant d'enrichir la communication ? Telles sont les questions posées dans notre recherche (cf. figure 2.5).

Nous tenterons d'apporter des éléments de réponse à ces questions en nous focalisant sur les caractéristiques de la communication réelle. Nous étu-

²Bernard Arcand écrit que : *le téléphone est un outil qui nuit beaucoup à la communication. Grâce au téléphone, on a de moins en moins besoin de se parler.* Il souhaite indiquer que les interactions entre les personnes s'amointrissent par ce support.

FIG. 2.5 – *La communication améliorée*

dierons différentes modélisations pour concenvoir un enrichissement de la communication.

2.3 Etude de la communication

La communication n'est pas une discipline académique, dans le sens où on désigne ainsi la physique ou l'économie, mais plutôt une discipline de carrefour où nombreux sont ceux qui y passent, mais peu qui s'y arrêtent.
[Willbur Schramm] - The Science of Human Communication, 1963

Le champ de recherche que nous présentons ici : la communication, ne fait pas partie de notre domaine de recherche. En revanche, nous étudierons et utiliserons à l'aide des nombreux écrits existants [24, 51, 90, 2] les différents modèles connus pour comprendre et faire progresser notre recherche.

Les niveaux de communication considérés

Nous avons présenté en page 25 les niveaux de communication sous la forme pyramidale. Il convient de définir plus précisément les types de communication que nous considérons comme importants dans notre recherche. Dans cette partie et par la suite, dans nos écrits, nous avons décidé d'étudier la communication dans le cadre de communication de type «chat» – interpersonnelle – dans les mondes virtuels coopératifs. Nous donnerons plus de précisions dans la partie réalité virtuelle sur les types de mondes et les interfaces utilisées. Nous avons porté notre attention sur la communication interpersonnelle car elle est à la base de toutes les communications humaines. Nous allons étudier plus en détail les spécificités de cette communication.

La communication interpersonnelle

Sans rentrer dans des détails philosophiques, nous poserons en premier lieu que toutes les «rencontres» entre individus, et où l'interaction entre les

individus est de mise, il y a communication interpersonnelle [45, 69, 63]. Nous avons écrit au début de ce chapitre que le moyen de communication le plus fréquent est la conversation de vive voix. Cependant, il existe de nombreuses autres formes de communication possibles [2] : l'écrit, les gestes, le regard, les expressions faciales, . . . Même le silence est un moyen de communication parfois bien plus efficace que d'autres plus bruyants. Nous voyons bien là qu'il n'est pas de notre ressort de classer toutes les formes. Dès lors, nous avons adopté une séparation simple en communication verbale et communication non-verbale. Nous allons maintenant étudier ces formes de communication pour en comprendre la substance.

La communication verbale

Nombreuses et variées sont les interactions faites par la voix. La conversation est un des fondements de la structure sociale humaine. La communication informelle par la parole occupe une place importante dans cette structure. La communication verbale montre un lien direct avec le contexte de l'information énoncé.

L'étude du langage permet d'observer et de comprendre comment les individus produisent et traitent l'information dans les échanges et la manière dont est employé le langage.

La communication non-verbale

Une grande partie des interactions sociales comprend des communications dites non-verbales. L'échange d'informations s'effectue aux moyens des gestes, du regard, . . . Certaines formes de la communication non-verbale permettent à deux personnes de communautés différentes de communiquer de façon universelle [26] : le rire est une expression universelle tout comme l'expression de la douleur.

Il faut aussi remarquer que la communication non-verbale n'est pas universelle, loin de là. Elle est souvent liée à un contexte culturel différent entre les individus. La couleur blanche signe de pureté peut être interprétée par certaines cultures comme signe de deuil (absence de vie).

Ces deux classes de communication sont aussi importantes l'une que l'autre. L'Homme a toujours inconsciemment utilisé des formes non-verbales pour communiquer. Limiter l'Homme à une seule de ces classes réduit l'expressivité des interlocuteurs qui devra dès lors s'employer à exprimer une forme sous une autre ou alors laisser place à plus de possibilités d'ambiguïtés. On comprend que la portée et les objectifs de ces modes de communication diffèrent entre elles.

Les autres études sur la communication

Les travaux sur la communications sont nombreux. Il nous est impossible de traiter toutes ces études. Il en existe un qui nous a semblé intéressé à citer : l'étude sur la distance de communication (la proxémique).

La proxémique : l'espace personnel existe même dans les mondes virtuels. Les spécialistes du domaine étudient le rôle joué par l'espace physique dans la communication entre les personnes.

La proxémique, dont les bases ont été jetées en 1963 par l'anthropologue Edward T. Hall (poursuivit par Burgoon [16]), pourrait varier selon les cultures. Mais de toute évidence la proxémique est un des nombreux contexte de la communication comme nous l'explique N. Yee [112] et ses portées sont importantes.

Pour une communication claire et réussie, les informations doivent de présenter sous différentes formes avec une concordance entre elles.

2.4 Les modèles de communication

Le processus de communication se réalise le plus communément par des interactions de trois éléments de base :

1. Un émetteur qui initie la communication.
2. Un canal qui permet l'envoi d'un message et d'informations.
3. Un récepteur qui reçoit l'objet de la communication.

Ce processus est plus ou moins complexe, suivant le nombre de personnes impliquées et les techniques employées comme moyen de diffusion. Il n'existe pas de modèle universellement accepté dans la recherche, pour représenter les phénomènes de la communication [51]. En effet, les modèles diffèrent selon qu'il s'agisse de la communication interindividuelle ou de masse, de la communication verbale, textuelle ou non-verbale – ex : gestes – mais aussi selon le mode de transmission. Nous présentons certains modèles afin de mieux comprendre les mécanismes mis en œuvre.

2.4.1 Le modèle de Shannon et Weaver

Le modèle de Shannon et Weaver [91] a été considéré pendant longtemps comme un modèle de base. Il est caractérisé par sa linéarité et sa simplicité.

Dans ce modèle schématisé figure 2.6, Une source d'informations – ex : intentions ou événements – transmet un message à un émetteur. Cet émetteur

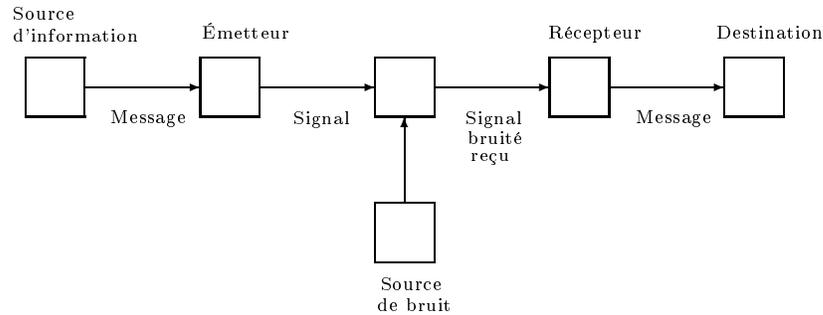


FIG. 2.6 – *Modèle communicationnel de Shannon et Weaver*

va transmettre un signal dans un environnement. Cet environnement subit les effets de différentes sources des bruits qui viennent altérer le signal original. Ce signal bruité est perçu par le récepteur qui le transforme en message avant de le transmettre au destinataire. C'est le modèle général de Shannon et Weaver d'un acte de communication.

Shannon³ et Weaver ont identifié trois niveaux de problème dans leur étude de la communication qu'ils ont énuméré sous forme de trois questions :

Question 1 *Avec quelle exactitude les symboles de la communication peuvent-ils être transmis ?*

Question 2 *Avec quelle précision les symboles transmis véhiculent-ils la signification désirée ?*

Question 3 *Avec quelle efficacité la signification reçue affecte-t-elle la conduite dans le sens désiré ?*

Ces trois questions, ces trois niveaux s'interpénètrent mais présentent néanmoins des caractéristiques qui les rendent indépendantes.

1. La question 1 est **technique**. Ce niveau est dans un sens le plus important car il influence tous les autres niveaux. Il met en avant le problème du transfert des séries de symboles depuis l'émetteur jusqu'au récepteur. En effet, ces symboles peuvent être détériorés et perdre leur sens initial.
2. La question 2 est **sémantique**. Celui-ci concerne le décodage du message qui, suivant la nature du décodeur, peut modifier le sens propre

³Shannon a travaillé pour la compagnie de téléphone Bell, il a surtout mené des études pour maximiser le signal. C'est pourquoi nous retrouvons dans le modèle 2.6 l'impact du bruit.

du message. Si l'encodeur et le décodeur sont d'«origine» différente, l'entrée de l'encodeur diffère de la sortie du décodeur.

3. Le question 3 concerne le niveau de l'**efficacité**. Toute communication tend à modifier un comportement et atteindre un objectif plus ou moins précis. Une communication «efficace» [91] est une communication qui à la capacité d'agir sur le récepteur.

Cette approche met en exergue de notre recherche les problèmes liés à la communication en général. Nous allons étudier ce même modèle avec l'introduction d'un monde virtuel qui médiatise les communications.

Spécificités aux mondes virtuels

Si on observe la place de la réalité virtuelle au sein de cette modélisation de la communication, on observe certaines spécificités (cf. figure 2.7) :

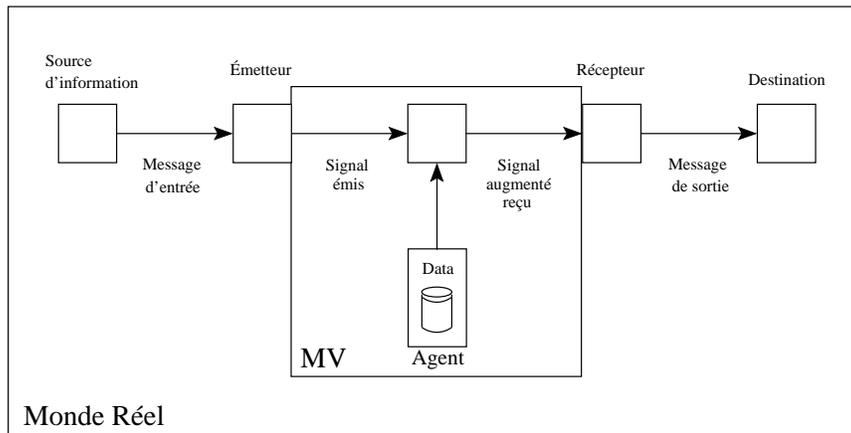


FIG. 2.7 – *Modèle communicationnel de Shannon-Weaver adapté à la RV*

- Le monde virtuel possède ses propres «bruits» tels que les mauvaises transmissions entraînant des coupures de connexions avec le monde virtuel simulé, rompant ainsi brutalement des actions en cours. Mais ces bruits pour la plupart proviennent d'une mauvaise utilisation du monde virtuel.
- Avec les différents outils à notre disposition, nous pouvons obtenir une concordance entre le signal envoyé et sa réception par le récepteur. Le monde virtuel (*MV*) offre une robustesse des messages transmis.
- En plus de pouvoir garder une cohérence entre le signal émis et le signal reçu, le *MV* peut aller au-delà et «améliorer» le signal envoyé par

le biais des informations disponibles.

- Un dernier constat que nous allons retrouver dans les autres modèles est le support commun. En effet, les actions s'effectuent au sein d'un monde virtuel, c'est-à-dire simulé, où toutes les informations du monde sont connues. Tout acte de communication peut être analysé et étudié lors de la transmission.

Conclusion

Ce modèle est une des bases de notre problématique. La communication réelle peut être endommagée par le «bruit», qui peut avoir pour conséquences de déformer la transmission ou faire disparaître le message. Le contexte extérieur est un élément qui s'ajoute au signal entre l'émetteur et le récepteur. Si ce bruit est introduit, le message contient des déformations et de ce fait, il peut provoquer des incertitudes.

A l'opposé de ces déformations, on trouve les mondes virtuels, même si ces mondes ne fournissent pas une aussi grande diversité des possibilités d'expressions, ils permettent de résoudre le problème de correspondance entre le signal reçu et le signal émis. De plus, nous verrons par la suite que contrairement au réel où l'on assiste à une diminution et une déformation des informations. Les *MV* peuvent améliorer les informations entre les utilisateurs par l'utilisation des informations détenues au sein de la simulation virtuelle.

2.4.2 Le modèle de Lasswell

Harold Dwight Lasswell est l'auteur d'un autre modèle descriptif de la communication [50]. Ce modèle très répandu a été proposé, en premier lieu, pour l'analyse de la communication de masse. Lasswell est connu pour sa phrase résumant son modèle : *Qui, dit quoi, par quel canal, à qui et avec quel effet ?*

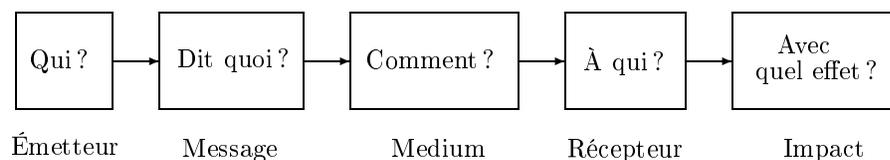


FIG. 2.8 – *Modèle communicationnel de Lasswell*

Dans le modèle figure 2.8, l'émetteur initialise tout le processus de com-

munication. Ce dernier veut transmettre un message. Pour arriver à ses fins, il adopte un médium pour la communication. Ce message contenu dans un médium est perçu par le récepteur qui réagit en conséquence. Cette réaction est l'impact de la communication.

Observons en détail les différentes parties de ce découpage, suivant le modèle de questions :

1. **Qui ?** : cette partie correspond au communicateur. Lasswell cite le journaliste de la presse, les radios et la télévision.
2. **Dit quoi ?** : sont les messages à proprement dit contenant une partie de l'information à transmettre.
3. **Par quel canal ?** : il représente l'autre partie de l'information : le support du message.
4. **À qui ?** : relatif au(x) récepteur(s) concerné(s), le public ou l'interlocuteur que le communicateur souhaite toucher.
5. **Avec quel effet ?** : désigne les impacts de la communication.

Par la suite, à l'aide de ce découpage simplifié de Lasswell, nos recherches ont pu être structurées. Il nous a permis de subdiviser le problème. Lasswell, porté par ce modèle, l'étend aux autres formes de communication et lie à chaque question une analyse particulière.

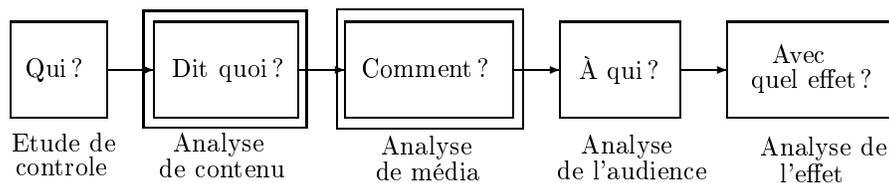


FIG. 2.9 – Études associées aux composantes de la communication

Cette énonciation des questions et des études (cf. figure 2.9), liées aux composantes de la communication, nous ont permis de mieux cerner les problèmes. En effet, l'«amélioration» de la communication correspond à une amélioration d'au moins une des composantes de la communication, soit une analyse du composant concerné. Nous pouvons dès lors nous concentrer sur la question :

Question 4 *Quelles sont les composantes pouvant être améliorées pour obtenir au final une communication améliorée ?*

Notre début de réponse a consisté à nous centrer sur une amélioration de la communication au niveau de l'«analyse de contenu» et l'«analyse de média» qui rentrent dans nos domaines de compétences. Lasswell suggère,

en outre, que la position de l'analyse du contenu est en amont de l'analyse des médias. Il lui apparaît important de traiter le contenu avant de traiter le support. Nous verrons que cette vision est aussi partagée par d'autres scientifiques de la communication tels que G. Gerbner.

Lasswell suggère la communication comme une forme d'outil de persuasion. Il convient de rappeler que ce psychologue et politologue suppose que toutes les communications ont pour but de modifier le ou les récepteurs et de provoquer un effet. Cette perception du problème relève d'une analyse différente de notre axe de recherche que nous n'approfondirons pas par la suite.

Conclusion

Lasswell a envisagé la transmission de l'information de manière linéaire dans le but d'agir sur le ou les récepteurs. Cette vision linéaire permet un traitement séquentiel de l'information d'où une utilisation de ce modèle dans les traitements d'informations. Le modèle de Lasswell nous a surtout permis de concevoir un découpage clair des composantes de la communication et d'en poser les analyses liées à celles-ci.

2.4.3 Le modèle de Gerbner

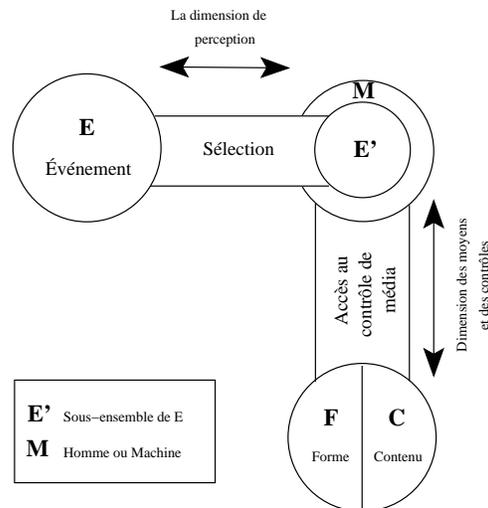
Un autre modèle nous est apparu intéressant à présenter et à analyser. Le modèle de George Gerbner, sociologue des années 1950, avait l'ambition de formuler un modèle général de la communication. Il présente en 1956 un modèle plus complexe que les précédents qui porte son nom [39]. Son modèle s'articule autour de deux propositions essentielles :

1. Le message est lié au contexte, à la «réalité» qui l'entoure, permettant de le renseigner sur sa signification.
2. Le processus de communication est un ensemble de deux dimensions : une dimension perceptive, une dimension pour le contrôle (cf. ci-après).

Le trait particulier de ce modèle est que l'on peut l'appliquer aux différentes formes de communication en fonction du contexte. Il convient à un acte de communication interpersonnel entre deux individus mais aussi au processus plus complexe de la communication de masse.

Pour comprendre les deux phases du modèle présenté figure 2.10, il est essentiel de comprendre que Gerbner adopte un point de vue «égocentrique» où le récepteur est au centre comme référentiel de toutes les actions.

La dimension perceptive : le processus de communication débute avec l'apparition d'un événement E , une modification de l'environnement. Cet

FIG. 2.10 – *Modèle communicationnel de Gerbner*

événement est perçu par une entité M représentant une personne ou machine perceptive (exemple : microphone ou caméra). Le point explicite dans ce modèle est la mise en relation de l'événement E et E' . Le receptrice ne peut saisir la totalité de l'événement E , il ne tient compte que de certaines parties de l'information perçues par ses « capteurs ». Nous pouvons alors noter que $E \subset E'$. Le processus de sélection est complexe, difficile à décrypter dans le cas d'une personne mais plus facile à décrypter dans le cas d'une machine.

La dimension de contrôle : Ce niveau concerne la transformation de la perception E' en signal concernant E – nommé CF . Ce signal se décompose en F – le signal, le contenant – et en C – le contenu. En effet, Gerbner admet qu'un contenu/une information peut être communiqué de manières différentes. Il suggère même d'offrir à l'utilisateur le choix de la forme F entre plusieurs formes possibles. Il apparaît que le choix de support F pour une événement E est une des plus grandes difficultés pour le communicateur.

Analyse

Au travers de ce modèle, Gerbner introduit plus clairement la notion de perception dans la communication. Dans le cas de la communication humaine, l'individu ne reçoit les événements que d'une manière machinale ou automatique. Tout événement est vécu par le prisme sélectif de nos sens. On retrouve, dans les travaux de ce sociologue de la communication, des explications [40] sur la perception des événements en rapport au contexte – le vécu, la culture et les idéologies – de la personne. Il nous met en garde contre le fait

que la communication introduit la notion de perception qui est un processus complexe où le contexte – les facteurs socio-culturels – sont importants.

Le contexte modifie notre perception du message. Il peut, dans certains cas, être une valeur négative et dérouter le message de son sens originel. Il peut, dans d'autres cas, donner des informations supplémentaires permettant une clarté du message. On peut alors se demander :

Question 5 *Le contexte est une source d'informations supplémentaires mais quand et comment peut-on l'utiliser afin qu'il soit un apport à la communication ?*

Gerbner ne s'arrête pas à cette constatation et continue son analyse en présentant l'ensemble de la forme F et du contenu C comme un tout. Un tout qu'il considère comme difficilement dissociable. Selon ce sociologue, même si nous arrivons à distinguer la forme du contenu, ces concepts sont intrinsèquement liés. F affecte la présentation de C . Le sens d'un message est la somme entre le forme et le contenu. Le choix de la forme suivant le contenu de l'information se révèle être aussi un aspect important de la communication.

Question 6 *Quelle forme choisir pour un message connu ?*

Spécificités de la réalité virtuelle

Le contexte dans le cadre d'un monde virtuel est identique pour tous les communicateurs. Un événement E créé par un monde virtuel peut être perçu sous différents aspects. Tout comme le réel, les différents sens et portées ne sont pas pleinement accessibles. Nous ne pouvons sélectionner qu'une partie – même dans le cas d'un monde virtuel – de l'événement E .

Nous pouvons aussi constater que la dimension perceptive est réduite. Les mondes virtuels ne permettent pas une transmission aussi riche que la communication de vis à vis. En contrepartie, un grand nombre des informations est directement accessible.

Conclusion

Ce modèle est très utile pour la compréhension du message, aussi bien au niveau individuel qu'au niveau social. Ce modèle est l'un des premiers à introduire le fait que toutes les entités sont des récepteurs. Il nous permettra par la suite d'illustrer des problèmes de communication et de perception. Il permettra aussi d'expliquer certaines parties des traitements de la transmission par les communicateurs.

2.4.4 Le modèle de Newcomb

En 1953, Theodore M. Newcomb présente le modèle ABX [70] triangulaire et introduit le rôle de la communication dans la relation sociale.

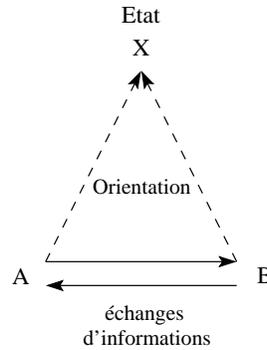


FIG. 2.11 – *Modèle communicationnel de Newcomb*

Pour Newcomb (cf. figure 2.11), A est le communicateur et B le destinataire. Il peut s'agir d'individus, mais aussi d'organisations. Dans cette relation, il introduit X l'environnement social qui peut représenter une personne ou une chose. Newcomb conclut ainsi qu'au travers de cette interdépendance, si A tend à changer, B et X changent à leur tour et inversement. En résumé, ce schéma nous montre que les échanges d'informations entre A et B tendent à rapprocher ces derniers vers un état de convergence (la situation commune X).

De manière plus pragmatique, la communication est pour ce chercheur, un moyen de permettre à deux ou plusieurs personnes de maintenir les orientations simultanées entre eux et à l'égard des objets de leur environnement extérieur. Il conçoit la communication comme un processus qui tend à maintenir une cohésion de la situation entre les trois éléments, par la transmission d'informations.

Question 7 *Le contexte de la communication ne permet-t-il pas de prédire certains actes de communication ?*

Conclusion

Ce modèle met en évidence le besoin de communication des individus afin de parvenir à une orientation commune nécessaire. On voit apparaître de cette manière la notion d'orientation de la communication.

2.4.5 Le modèle de Westley et Mac Lean

Le modèle décrit par Westley et Mac Lean [108] est une reprise du modèle de Newcomb avec une extension à la communication de masse qui possèdent ses caractéristiques propres. Il est intéressant d'étudier ce modèle qui est une fusion du modèle de Shannon, de Newcomb et de Gerbner.

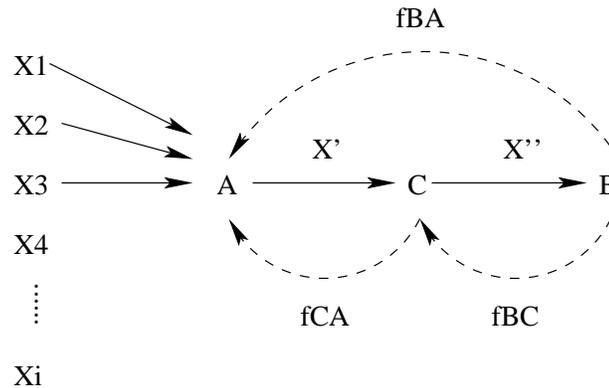


FIG. 2.12 – *Modèle communicationnel de Westley et Mac Lean*

Analyse

En premier lieu, présentons les différents composants du modèle schématisé figure 2.12 :

- X_i est un événement extérieur dans le contexte.
- A est un individu qui souhaite communiquer des informations concernant l'événement X .
- C est le média ayant pour objectif de transmettre le mieux possible les informations.
- B est le récepteur qui attend des informations sur X .
- X' est une partie des informations de A choisie par C .
- X'' est le message transmis modifié au préalable.
- les fonctions fXY sont les effets de retour.

Sans rentrer dans les méandres du modèle, tout comme Gerbner, Westley et Mac Lean estiment qu'une partie des informations de l'environnement X est sélectionné par le communicateur A qui à son tour, va transmettre l'information avec une certaine modification. Le fait est que la communication possède comme « agents-clés » : A et C . Ils ont un rôle actif par rapport à B qui est dépendant de ces entités. Toutes améliorations de la communication devraient donc être liées à l'entité A afin de lui donner la possibilité d'offrir un maximum d'informations de l'événement X – qui se traduit par une

amélioration des informations transmises par A – et une amélioration des choix de l'entité C – qui se traduit par une compréhension des informations disponibles et une transmission adaptée à l'information.

Question 8 *Quand et comment les «agents-clés» peuvent-ils aider la communication à être plus pertinente ?*

Conclusion

Ce modèle regroupe de nombreuses informations et offre une certaine fusion des autres modèles. Westley-MacLean écrivent que les médias servent à satisfaire des besoins d'information, le besoin de communication des communicateurs. Ils citent l'intérêt des médias – les mondes virtuels dans notre contexte – qui est d'offrir une vision partagée sur l'environnement. Ils soulignent de ce fait la notion de «perceptions partagées» au sein d'une entité commune.

2.4.6 Le modèle de Riley et Riley

Les premiers modèles de communication négligient l'environnement dans lequel l'acte se déroulait. Cela est dû en grande partie aux expériences en laboratoire où les conditions expérimentales diffèrent des conditions réelles.

Riley et Riley [82] ont cherché à formuler un modèle tenant compte de l'environnement dans le processus de communication. Ils affirment que la communication de masse est un système social parmi d'autres dans la société.

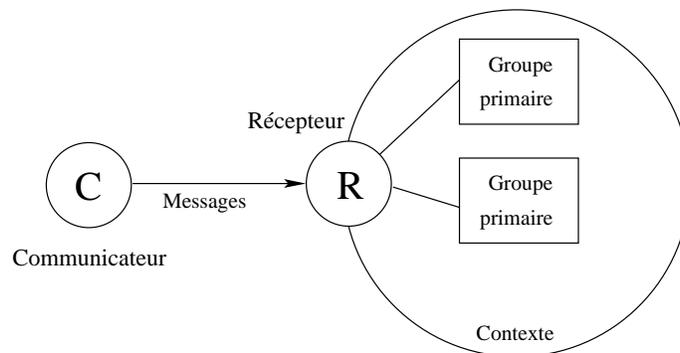


FIG. 2.13 – *Modèle communicationnel sur l'émission d'un message*

Ils se basent sur le modèle traditionnel de la communication de stimulus et de réponses. Ils pensent que les communicateurs ont toujours l'intention

d'influencer le ou les récepteurs au travers des messages particuliers (cf. figure 2.13). Les individus, disent-ils, même lorsqu'ils sont seuls au moment de recevoir les messages, ne sont pas des atomes isolés, mais font partie des groupes primaires (famille, groupe de travail par exemple) avec lesquels ils partagent un certain nombre de comportements et de valeurs, et des groupes de référence, qui leur servent de guide dans les décisions. Par groupe de références, on entend le groupe qui aide l'individu à définir ses attitudes et ses valeurs. L'individu – émetteur ou récepteur – est influencée par le groupe auquel il est le plus rattaché. Le communication peut ainsi être influencée non seulement pour la sélection mais aussi la perception des messages.

Question 9 *Comment se définissent les groupes d'influences qui permettent de comprendre les communications ?*

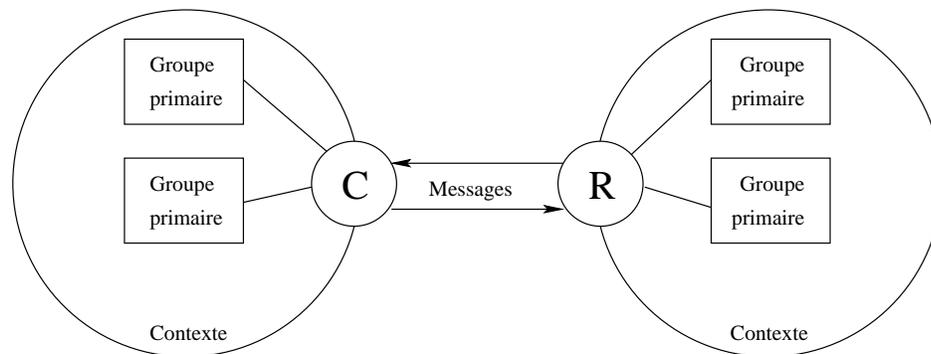


FIG. 2.14 – *Modèle communicationnel de Riley et Riley*

En bref, le modèle des Rileys (cf. figure 2.14) fait apparaître la liaison complexe entre communication de masse comme processus social particulier et l'ensemble de la société.

Spécificités à la réalité virtuelle

Dans un environnement virtuel, nous avons vu qu'une partie du contexte est commun à tout le monde. A travers ce modèle, nous voyons apparaître explicitement la notion de compétences et l'importance de ces groupes de compétences.

conclusion

Nous voyons apparaître dans ce modèle la notion de «contexte prédéfini» où chaque utilisateur possède un contexte défini par la tâche qu'il souhaite

effectuer ou la tâche qui lui a été attribuée. L'utilisateur ainsi contextualisé fournit ou retire des orientations connues.

2.5 Conclusion

La communication : omniprésente, mais toujours imparfaite.
[J.F. Dortier, dictionnaire des sciences humaines, 2004, introduction]

La communication limpide et transparente est un mythe. Les messages sont souvent ambivalents, le récepteur sélectionne les données. Les véritables enjeux et les niveaux impliqués sont souvent cachés.

2.5.1 Les problèmes de communication

En nous appuyant sur cette étude, nous pouvons distinguer les cas où la communication pose problème. Étudions le cas général dans le domaine de la psychologie. Au sens large du terme, les troubles de la communication englobent toutes les difficultés à :

1. Capter l'information ;
2. Traiter l'information reçue (sous toutes les formes possibles) ;
3. Connaître/apprendre un système de symboles ;
4. Stocker/accéder à l'information apprise ;
5. Formuler un message ;
6. Transmettre l'information qui correspond à nos intentions de communication.

On réalise que toutes ces conditions sont nécessaires pour une communication. Le point 1 exprime l'idée d'être en mesure de pouvoir capter l'information. Une personne derrière un mur insonorisé a des difficultés à communiquer. Si l'interface d'émission est incompatible par rapport à l'interface de réception, il nous est impossible de traiter l'information. Pour communiquer, il faut aussi comprendre le message (en retirer une information). L'entité doit être en mesure de connaître un système de symboles afin que les informations puissent être traitées. Ceci implique un stockage et un accès à l'information du système de symboles. Les deux derniers points mettent en avant le fait d'avoir tous les moyens de transmission d'un message.

2.5.2 Le récepteur

Dans la communication, le plus compliqué n'est ni le message, ni la technique, mais le récepteur. [Dominique Wolton]

Le récepteur n'est pas un élément passif de la communication. Lors de la lecture, l'écoute ou la visualisation d'un message, l'attention du récepteur est toujours sélective. Les informations contenues dans un message sont toujours filtrées par divers processus de sélection. Ce que l'émetteur veut transmettre n'est jamais parfaitement compris par le récepteur. Le récepteur filtre, décode, sélectionne, réinterprète les informations reçues. Il « traite l'information ». Ce traitement sort du cadre de notre recherche. Nous laisserons de côté le traitement par le récepteur qui concerne le domaine de la perception, de la mémorisation et plus généralement de la psychologie cognitive.

2.5.3 La forme agit sur le contenu

Les meilleurs conférenciers ne sont pas toujours les plus grands penseurs. Inversement, les plus médiocres orateurs ne sont pas forcément ceux qui ont les plus mauvaises idées. [J.F. Dortier]

Une des leçons des théories de la communication nous informe que « la forme du message rejaillit sur le contenu ». Un même message sera perçu différemment suivant sa forme et donc provoquera des effets différents. Dès lors la maîtrise du contenant permet une meilleure compréhension du contenu. L'amélioration du contenu est intimement liée aux possibilités du contenant. Une information peut être noyée si elle s'exprime sous une forme complexe ou impropre.

2.5.4 Les lois de la bonne communication

L'incommunicabilité ? Ce n'est pas qu'on ne communique pas assez. C'est que l'on communique trop et mal. [Robert Lalonde]

Une bonne communication tient compte de trois facteurs :

1. La clarté du message : avec un minimum d'ambiguïtés
2. La prise en compte du récepteur : étude de son contexte, sa vision, ses intérêts.
3. La qualité de la relation établie : la communication est pertinente, fournir des informations non ambiguës.

Suivant les deux premiers points, nous avons axé nos recherches. Ils seront traités de manière plus approfondie par la suite. Le dernier point est plus un constat de résultats qu'une méthode permettant une amélioration de la communication. En effet, une qualité de communication ou de la relation établie n'est quantifiable dans notre approche qu'au travers d'une évaluation de la pertinence. Nous approfondirons ce point par la suite.

2.5.5 Conclusion

La chose la plus importante en communication, c'est d'entendre ce qui n'est pas dit. [Peter Druker]

Cette étude sur la communication a permis de soulever de nombreuses interrogations. Nous allons dans la suite de nos écrits essayer de répondre à certaines d'entre elles. Nous allons à présent étudier les différents outils qui nous ont permis d'aborder ces questions.

3 La réalité virtuelle

Sommaire

3.1	Introduction	46
3.1.1	Dispositifs immersifs	47
3.1.2	Dispositifs semi-immersifs	47
3.1.3	Dispositifs non-immersifs	48
3.2	Applications	49
3.3	Apports et limites des mondes virtuels	49
3.3.1	Sequence World : visualisation d'indices	50
3.3.2	Newton World : l'influence de la gravité	51
3.3.3	Limitations des mondes virtuels	51
3.4	Étude des communications dans les mondes virtuels	53
3.4.1	Introduction de l'étude	55
3.4.2	Les Environnements Virtuels	55
3.4.3	Dive : Distributed Interactive Virtual Environment	55
3.4.4	3DWebmaster : le langage SCL	58
3.4.5	Ulysse : agent conversationnel interactif	60
3.4.6	VREng : Virtual Reality Engine	63
3.4.7	Diamond Park : système social virtuel	64
3.4.8	Active World : l'un des premiers EV	67
3.4.9	Open Croquet : mondes virtuels collaboratifs	69
3.4.10	Habbo hotel : un monde en ligne	71
3.4.11	Spin-3D : pour un travail collaboratif	73
3.4.12	LEA : LIMSI Embodied Agent	75
3.4.13	Divers mondes virtuels	77
3.4.14	Synthèse de l'étude	81
3.4.15	La RV et la communication	84
3.5	Conclusion	88

The physical world, the thing on the other side of your sense organs, is received through these five holes, the eyes, and the ears, and the nose, and the mouth, and the skin. They're not holes, actually, and there are many

more senses than five but that's the old model, so we'll just stick with it for now. [Jaron Lanier]

Dans ce chapitre nous allons d'abord présenter notre outil de recherche : la «Réalité Virtuelle» (*RV*) et plus précisément les «Environnements Virtuels» (*EV*). Ces derniers permettent de simuler un ou plusieurs «Mondes Virtuels» (*MV*). Après une introduction à la réalité virtuelle et sur quelques dispositifs existants, nous présenterons des réalisations qui ont su exploiter au mieux les possibilités de la réalité virtuelle. Ensuite, nous étudierons les communications au sein d'environnements virtuels existants. Finalement, nous conclurons sur les réponses de la réalité virtuelle aux questions de la communication.

3.1 Introduction

Depuis quelques années, la réalité virtuelle est entrée dans nos coutumes et dans notre langage. Ce que l'on croyait encore être dans le domaine du rêve et de la science-fiction, progresse et arrive à notre portée. Il n'est plus rare maintenant de se donner rendez-vous sur un site virtuel ou un réseau pour dialoguer et se rencontrer. La littérature et le cinéma ont été un élément important dans l'introduction de la *RV* dans nos mœurs. Nombreuses sont les recherches en cours et à venir. Mais quelles ont été les leçons tirées de ces travaux ? On y retrouve souvent une nécessité d'enrichir une approche purement technique. Les objectifs sont pour la plupart de créer et d'assister le comportement humain dans les actes d'interaction avec un monde virtuel.

Lorsque nous évoquons la réalité virtuelle, nous nous imaginons immergés dans un décor irréel bardé de capteurs, d'un casque ou encore de «cyber-instruments» – ex : cybergloves. Mais ce n'est qu'une vision réductrice de la réalité virtuelle.

Qu'est ce que la réalité virtuelle ? On parle de réalité virtuelle lorsque nous sommes dans un univers de synthèse où l'on peut éprouver des sensations comparables à la réalité, et où l'on peut agir avec différents composants de cet univers. C'est avant tout un système informatique visant à créer des environnements qui simulent un monde réaliste ou un monde fantastique (cf. figure 3.1) donnant ou non ainsi à l'utilisateur l'impression de réalité. D'où le sens de la phrase de Jaron Lanier mise en exergue.

Il existe de nombreux systèmes de réalité virtuelle que nous pouvons classer en trois types reconnus par les scientifiques et que nous allons présenter de façon concise.



Karobas : un monde réaliste



World of Warcraft : un monde fantastique

FIG. 3.1 – Exemples de mondes virtuel

3.1.1 Dispositifs immersifs

Les CAVEs en sont un exemple idéal (cf. figure 3.2). L'utilisateur est plongé au milieu d'une scène projetée autour de lui à l'aide de six écrans couvrant tout son champ de vision. L'impression d'immersion est alors importante (cf. rapport sur la plate-forme PerfRV[11]). Ces systèmes immersifs comportent des interfaces tels les casques stéréos ou encore les gants sensitifs à retour d'efforts (gants haptiques).

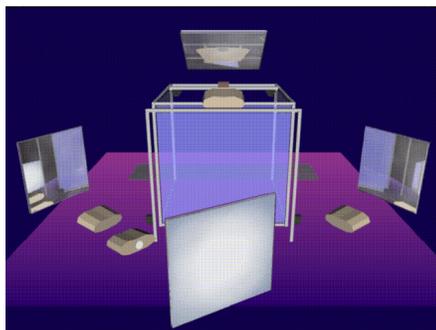


FIG. 3.2 – Dispositif immersif : un CAVE

3.1.2 Dispositifs semi-immersifs

Les «Tables de travail»(cf. figure 3.3) («WorkBenchs»), ou encore appelés demi-CAVEs, offrent une large vision de la scène mais le support ne recouvre pas tout le champ de vision de l'utilisateur. En effet, ces tables ne couvrent pas tout le champ de visualisation[25]. Néanmoins, elles sont présentées comme des outils plus accessibles que les systèmes immersifs par

leurs tailles réduites et leurs coûts.



FIG. 3.3 – *Dispositif semi-immersif : un Workbench*

3.1.3 Dispositifs non-immersifs

Les systèmes à l'écran ont des interfaces plus limitées : le champ visuel occupé par l'environnement virtuel coïncide avec l'écran d'un moniteur. Il est possible d'améliorer cette vision par le port de lunettes anaglyphes¹ ou encore stéréoscopiques², procurant une impression de relief. Ces systèmes peuvent également comporter des «cyber-instruments» comme des gants dotés de capteurs de position, représentés dans l'environnement et visibles à l'écran. Nous pouvons voir figure 3.4, un dispositif initialement immersif utilisé pour améliorer le dispositif non-immersif – le phantom, un stylet à retour d'efforts. Ces instruments rendent plus naturelles certaines actions spécifiques telles que déplacer, actionner, agir sur les éléments du monde virtuel.

Actuellement, ces dispositifs non-immersifs sont les plus répandus par leurs emplois, les dispositifs semi-immersifs et immersifs ne percent que faiblement dans le domaine du grand public et restent des dispositifs exclusivement de recherche.

¹Un anaglyphe est une image imprimée pour être vue en relief, à l'aide de deux filtres de couleurs différentes (lunettes anaglyphes).

²Lunettes occultants alternativement les yeux pour permettre l'affichage d'une image différente pour chaque œil.

FIG. 3.4 – *Dispositif non-immersif : un phantom*

3.2 Applications

Les systèmes de réalité virtuelle ont des applications dans de multiples domaines. Dans les domaines militaires et de l'aéronautique, on utilise surtout des équipements spécialisés pour exécuter des séquences d'entraînements permettant un apprentissage plus proche du réel – ex : simulations de vol. Ils constituent des secteurs importants en matière de financement et de ressources. Dans le domaine de l'architecture ou de la conception assistée par ordinateur (*CAO*), la réalité virtuelle permet une visualisation directe du résultat final. Par exemple, il est tout à fait possible de circuler au milieu de structures 3D virtuelles, représentant des bâtiments en construction. Actuellement, les réalisations de pointe en réalité virtuelle sont en médecine. A ce jour, il est possible d'assister à la réalisation de simulations d'opérations collaboratives avec des utilisateurs géographiquement éloignés. Un dernier domaine est l'art. C'est un domaine dans lequel on cherche à aller au-delà des lois de la Physique (cf. figure 3.5).

3.3 Apports et limites des mondes virtuels

Les *EV* ont leurs propres spécificités en matière d'interaction. Ils permettent de simuler un ou plusieurs mondes virtuels et introduisent de nouvelles possibilités d'interactions et avec elles de nouvelles limites. Afin de comprendre ces possibilités des mondes virtuels vis-à-vis du réel, nous allons présenter des programmes qui permettent un aperçu de notre intérêt suscité par le *MV*.

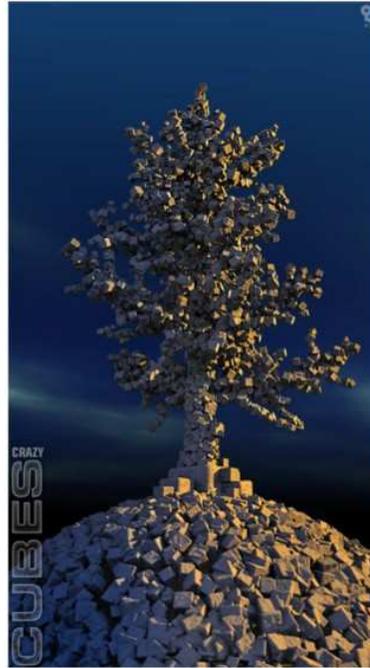


FIG. 3.5 – *Un arbre de cubes*

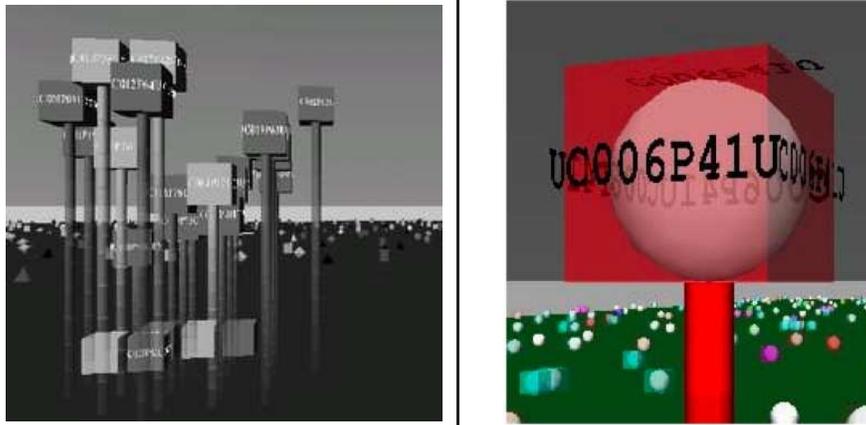
3.3.1 Sequence World : visualisation d'indices

«Sequence World»[83] est une application en cours de développement permettant une visualisation des colossales bases de données de la *GenBank*³. Le problème majeur de cette base est sa taille : les données abstraites existent mais à cause des grandes quantités d'informations numériques, la lecture et le décryptage des informations contenues dans ces données restent une difficulté. Seule une analyse approfondie permettrait de révéler des informations utiles dans cette masse indescriptible d'informations. Afin de palier à ce problème, un programme a été conçu pour permettre une visualisation aisée et une mise en évidence des éléments suivant une méthode de classification.

Sur la figure 3.6, nous pouvons observer un ensemble de sphères entourées par des cubes et le tout monté sur des cylindres. En regardant de manière plus détaillée, on y voit des jeux de transparence entre le cube et la sphère qui la contient. Ces effets sont en réalité des indications sur des indices de confiance. Il nous est dès lors possible de situer d'un seul regard les éléments pertinents.

Ce programme présente plusieurs intérêts, les informations visibles sous la

³Banque générale américaine sur les séquences de protéines et des séquences d'ADN

FIG. 3.6 – Vue sur un monde de *SequenceWorld*

forme d'un tableau à deux dimensions sont transformées sous les traits d'un monde en trois dimensions. Cette troisième dimension, difficilement représentable dans le réel, permet une nouvelle mise en forme de l'information. Cet enrichissement de la visualisation des données améliore la perception d'informations.

Les mondes virtuels permettent une modélisation et une visualisation d'informations sous divers formats. Nous pouvons citer les travaux de Silva et al. [93] où l'*EV* est utilisé comme une source d'enrichissement pour la présentation d'informations suivant un format adapté aux besoins.

3.3.2 Newton World : l'influence de la gravité

«Newton World»[86] est l'un des premiers mondes virtuels créé dans le but de comprendre l'impact de la gravité dans le monde réel. Il fournit un outil de visualisation des effets de la gravité sur les objets. Ces derniers ont différentes masses et la manipulation rapide des différents «variables» du monde, nous fournit une appréciation du rapport masse/gravité.

3.3.3 Limitations des mondes virtuels

On observe des différences entre nos expériences avec les mondes virtuels et avec le monde réel. Nous considérons souvent ces différences comme étant dues à des problèmes techniques. L'objectif des *MV* n'est pas seulement de rapprocher le monde virtuel du monde réel comme c'est le cas pour les simulateurs. Ils permettent un apprentissage de la réalité.

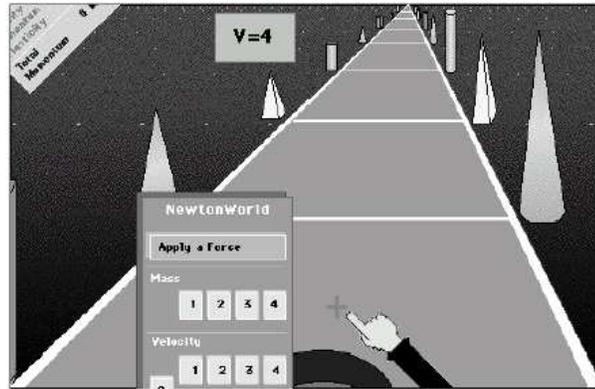


FIG. 3.7 – Vue sur un monde de Newton World

On retrouve une partie des apports et des limitations dans l'état de l'art de C. Chaillou[18] où il décrit la possibilité de modifier et de rendre à l'échelle humaine des visualisations difficiles comme la visualisation au niveau moléculaire. Dans notre exemple de Sequence World (cf. figure 3.6), nous pouvons visualiser des interactions entre protéines.

Dans les travaux de Mike Fraser[34], nous voyons une méthode de mise en évidence de certaines différences entre ce qui nous semble être perçu et ce qui est réellement perçu. Ses travaux mettent en exergue les limitations techniques du *MV*. L'illustration la plus flagrante est le cas du champ de vision.

En effet, les limitations techniques (l'écran) nous empêchent de voir le monde avec notre champ de vision maximale. Or, cette limitation n'est pas visible pour les autres utilisateurs. Dans la vie courante, il est possible d'estimer la vision à l'aide de la direction de la tête et du regard. Mais lorsqu'il s'agit de *MV*, le problème est tout autre.

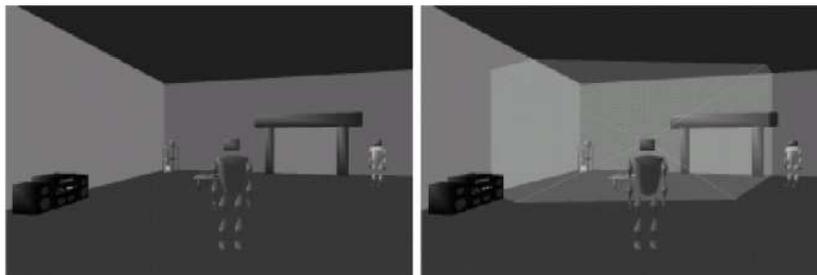


FIG. 3.8 – Illustrations d'une limitation des *MV*

Dans cette illustration (cf. figure 3.8), l'avatar (la représentation virtuelle

de l'utilisateur dans le monde virtuel) de droite fait des signes gestuels à celui situé au premier plan. Ce dernier ne répond pas aux appels et on s'interroge : «Pourquoi?». La réponse nous est donnée par l'affichage de manière surréaliste les limites du champs de vision : l'avatar de droite n'est pas visible pour celui-ci. Les appels gestuels ne peuvent être perçus par l'utilisateur. Ainsi, si l'avatar de droite souhaite être connu de l'utilisateur dont l'avatar nous apparaît au premier plan, il doit s'avancer.

Dans l'article de Choi et al.[20], les auteurs proposent des interactions en 3D de manière simplifiées. Leurs travaux permettent de résoudre un des problèmes de la manipulation interactive : l'imprécision. De nombreux cas usuels mettent en avant cette imprécision. Le simple fait de déposer une fleur dans un vase relève de la prouesse avec les outils standards d'interaction. Nous pourrions discuter des problèmes sur l'interpénétration et les impacts sur certains objets en appliquant un algorithme de détection de topologie spatiale[74, 78]. Retenons que certaines interactions simples dans le réel deviennent une véritable épreuve d'habileté dans les mondes virtuels. Il apparaît donc important de faire attention à ne pas compliquer certaines tâches.

Conclusion

La réalité virtuelle est une source d'enrichissements de la réalité. Bien employée, elle se révèle être un outil performant et efficace. Elle possède des limitations qu'il faut comprendre et tenir compte pour la mise en place d'un système. Dès lors, nous avons cherché les apports possibles et les limitations existantes dans différents *EV* du point de vue de la communication.

3.4 Étude des communications dans les mondes virtuels

Nous étudions dans cette partie les différentes plate-formes de réalité virtuelle intéressantes dans le cadre de notre recherche, c'est-à-dire pour ce qui concerne la communication. Nous allons comparer les différentes interactions présentes au sein de ces mondes. Dans ce vaste domaine de la réalité virtuelle et des environnements virtuels, nous avons orienté notre étude sur la communication en majeure partie vers des dispositifs non-immersifs dits classiques composés pour la plupart de deux supports de communication : le clavier et la souris. Ils représentent la grande majorité des *EV* actuels.

La représentation de mondes réels ou imaginaires s'est imposée comme une nécessité dans de nombreux domaines. Elle permet une étude approfondie

die des phénomènes antérieurs et postérieurs à un événement. Nous pouvons citer les travaux en mécanique des fluides[38] qui permettent une visualisation des turbulences. L'apport des *MV* a été de pouvoir visualiser des informations sous une forme plus «humaine»⁴. Mais une bonne représentation n'est pas suffisante. Pour utiliser pleinement les possibilités offertes par la réalité virtuelle, il nous faut interagir avec ce monde afin d'en soutirer des informations pertinentes.

Les méthodes de communication sont nombreuses. Dans bien des cas, c'est la nature de l'application qui dicte les composantes nécessaires à l'élaboration de ces plate-formes. Les modes/méthodes de communication dans les souvent mondes virtuels demeurent mal ou peu exploitées. Nous présentons une étude bibliographique sur les différents environnements virtuels existants ainsi qu'une analyse des différentes méthodes de communication. A la suite de l'approfondissement de la communication que nous avons mené au cours du chapitre 2, nous étudions les différents types de communication présents dans les *EV*.

⁴*L'homme est la mesure de toute chose.* [Protagoras]

3.4.1 Introduction de l'étude

Cette étude est un regroupement des connaissances acquises au cours de nos observations et utilisation des différents *MV*. Ces écrits sont orientés vers la communication entre la machine et son utilisateur. Dans le cadre de notre recherche, nous allons étudier la communication dans les mondes virtuels sans pour autant explorer plus en avant les sens complexes de la communication virtualisée (cf. A.Grumbach[42]).

Cet état de l'art porte sur une étude comparative des diverses méthodes de communication présentes dans les *EV*. Cette analyse est un dépouillement des différentes spécificités de la communication. Pour chaque environnement, nous avons tenté de comprendre les communications interpersonnelles et individuelles dans le cadre de notre recherche.

Méthode d'analyse

Après avoir sélectionné une liste d'environnements susceptibles de nous aider dans notre démarche. Dans le cadre de nos recherches, notre choix s'est porté avant tout sur des mondes virtuels non-immersifs avec comme support de communication des outils standards : des claviers ou des souris pour la plupart. L'analyse s'est effectuée à l'aide du modèle de Lasswell[50] (cf. chapitre 2 page 32) pour extraire et présenter les communications intéressantes. De cette analyse, nous avons obtenu un rapport technique sur la communications des mondes virtuels[101].

Axe de lecture : Nous allons présenter différents points intéressants pour chacun des *EV*. Une lecture possible est de se rendre directement à la synthèse (cf. Section 3.4.14 page 81) pour ensuite revenir aux différents *EVs*.

3.4.2 Les Environnements Virtuels

3.4.3 Dive : Distributed Interactive Virtual Environment

Dive est une plate-forme de réalité virtuelle multi-utilisateurs, où les participants peuvent naviguer dans un univers 3D. Ils peuvent se promener, rencontrer et interagir sur les différents éléments du monde. Les recherches et applications de *Dive* [36, 35, 37] sont nombreuses, citons notamment les interactions multi-modales ou encore le contrôle de robot à distance.

Notre choix s'est porté sur *DIVE*[37]⁵ car c'est un système complet ayant

⁵Source disponible sur : <http://www.sics.se/dive>



FIG. 3.9 – Vues du monde de DIVE

des objectifs multiples.

La communication inter-utilisateurs

Dans cette plate-forme, il existe plusieurs possibilités de communication. La première est assez triviale et connue, elle permet la communication avec d'autres utilisateurs. Cette communication s'effectue de manière dite «classique» : nous ouvrons un menu où la liste des utilisateurs présents dans le monde est affichée, nous y choisissons un utilisateur que nous souhaitons contacter. Cette méthode nous donne l'ensemble de tous les utilisateurs, même ceux qui ne sont pas directement visibles. Nous pouvons ainsi sélectionner une personne non visible.

Il persiste néanmoins un problème dans cette méthode qui peut s'apparenter un client IRC⁶(cf. Figure 3.10) : pour converser avec une personne seule, nous devons connaître son pseudonyme. Seule cette connaissance permet de faire la correspondance entre le dialogue de l'utilisateur et son avatar virtuel. Or si nous n'avons pas connaissance du pseudonyme, il est difficile, voire impossible de communiquer avec celui-ci. Il est possible de l'observer

⁶IRC : «Internet Relay Chat», protocole permettant de dialoguer de manière textuelle au sein de canaux.

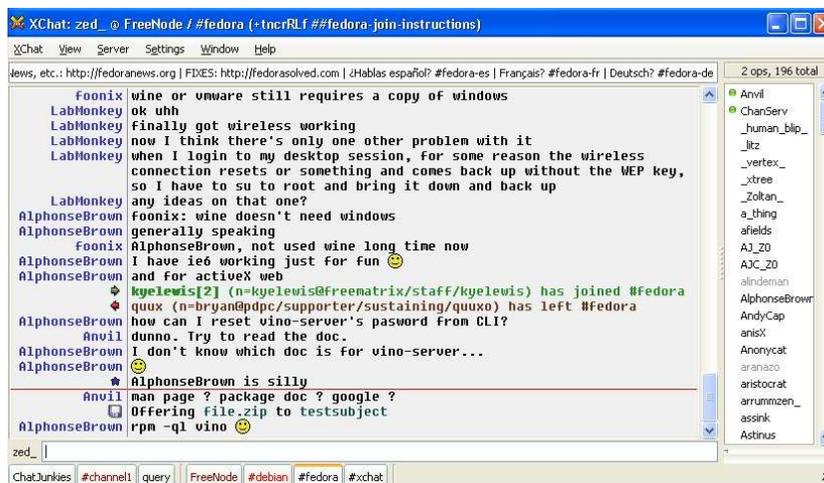


FIG. 3.10 – Image d'une fenêtre de XChat

et de le désigner (par pointage) mais là encore, certaines situations rendent impossible cette action (ensemble d'utilisateurs regroupés). Concernant le dialogue proprement dit, il se déroule le plus souvent sous la forme d'un «Chat» classique. Lorsqu'un utilisateur souhaite transmettre un message, il envoie son message dans un «canal» où tous les utilisateurs du canal voient le message en question. Il reste à faire, la correspondance entre le message envoyé et l'utilisateur correspondant.

La communication avec les automates

La seconde possibilité est une communication avec les éléments du monde, tels que les robots ou les agents. Dans le monde d'introduction de *Dive*, nous pouvons interagir avec un automate. Dès que nous nous approchons de ce dernier, il perçoit notre présence, se tourne vers nous et nous salue. Le *MV* de *Dive* envoie toutes les informations concernant l'utilisateur, permettant à l'automate de connaître la position ou nous saluer à notre approche.

L'interaction avec le monde

Le troisième type de communication possible, qui est plus une interaction qu'une forme de communication, est celui avec le monde lui-même. Nous pouvons modifier et transformer le monde à notre guise. Les restrictions concernant la transformation du monde sont inexistantes. L'interaction est illimitée avec le monde, elle s'effectue le plus souvent à l'aide de fenêtres et de menus, où les différents paramètres sont modifiables.

Analyse des types de communication

Lors de la communication avec d'autres utilisateurs, tous les messages et informations échangés sont sous la forme de textes. C'est l'utilisateur qui travaille et interprète les messages qu'il reçoit. C'est une communication restrictive. Elle n'utilise pas les possibilités de la réalité virtuelle. La fenêtre de dialogue apparaît à l'écran (occultant une partie de notre vision) et nous ne voyons pas nécessairement la personne, et même si elle était devant nous la communication tend à rester au niveau de la fenêtre. Il apparaît alors clairement que nous pourrions à l'aide de bulles affichées au-dessus des personnes, voir les interlocuteurs.

Un dernier type de communication est présent entre les utilisateurs : le pointage. On communique ainsi notre point d'intérêt aux personnes pouvant observer le vecteur rayon de désignation. Cette désignation permet de transmettre indifféremment le désir de montrer un lieu ou de sélectionner un objet. Nous avons alors une ambiguïté sur l'impact recherché qui peut être source d'erreurs.

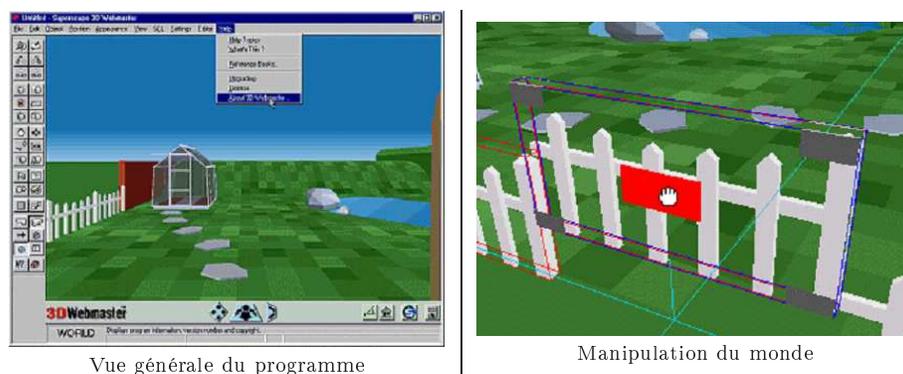
Conclusion

Il existe différentes causes pour lesquelles une communication s'effectue dans *Dive*. C'est un monde généraliste avec des objectifs divers, il est un support pour les différentes recherches en cours. On peut l'opposer à un monde professionnel possédant un objectif précis ou des tâches précises.

Le souhait de déclencher une communication au sein de cet *EV* provient d'une envie de rencontres avec les personnes présentes. Dans ce cas, l'utilisation de la boîte de dialogue s'avère peu efficace pour les raisons évoquées plus haut. Les autres types de communication peuvent parfois conduire à des ambiguïtés telle que la désignation ou encore une mauvaise interprétation à la réaction des automates.

3.4.4 3DWebmaster : le langage SCL

Superscape - *3DWebmaster*[107] est un logiciel commercial permettant une création rapide et une manipulation aisée de mondes virtuels. Il n'est plus actuellement supporté par ses créateurs mais possède une caractéristique qui le rend intéressant et qui justifie sa présence dans cette étude. Il s'appuie sur un langage propre (*SCL*) qui permet de spécifier des comportements pour différents objets du monde et offre ainsi une voie de communication nouvelle.

FIG. 3.11 – Vues de *SuperScape-3D Webmaster*

C'est en partie grâce à ce langage que nous pouvons effectuer de nombreuses actions au sein du *MV*.

Analyse des types de communication

Les différents émetteurs ne sont pas aussi nombreux que dans les plateformes de recherches tel que *Dive* (cf. Section 3.4.3 page 55) ou *VREng* (cf. Section 3.4.6 page 63). L'ensemble des possibilités de communication se limite à une communication entre le monde et l'utilisateur, le logiciel est mono-utilisateur – du moins dans la version que nous connaissons. Il n'est donc pas possible de rencontrer d'autres avatars avec lesquels nous pouvons dialoguer ou interagir.

La communication entre entités virtuelles

La communication qui nous intéresse dans ce logiciel est celle qui existe entre les objets. En effet, les objets communiquent entre eux mais aussi avec le monde et l'utilisateur. Tout passe par l'emploi du langage *SCL* qui a été créé dans cette optique afin de formaliser les actions possibles. Les entités peuvent ainsi utiliser certaines informations du monde pour effectuer une action. A titre d'exemple, il est possible de donner à un objet la propriété de disparaître lorsque nous nous approchons de ce dernier. *SCL* offre une nouvelle manière de concevoir et de manipuler les mondes. Il permet de concevoir un monde réactif et de donner un semblant d'«ambiance».

La lecture d'un programme *SCL* nous montre la transmission d'information et parfois même l'action entre les objets. Ils peuvent fournir à d'autres entités les informations tel que le clic de l'utilisateur rendant visible ou non une autre entité.

La communication entre les objets est d'une nature différente. En effet, les messages transmis entre les objets *SCL* et les autres sont des signaux. En effet, dès qu'une action, par exemple un clic, est effectuée sur un objet, l'objet peut interpréter ce clic et générer un autre signal. Cette communication par signal est interactive dû en grande partie au langage *SCL*.

Les supports de communication

Il existe plusieurs supports possibles pour communiquer avec le monde ainsi simulé. D'abord, une communication visuelle et sonore : avec une utilisation de la souris pour dessiner et pour manipuler les propriétés. Un emploi du clavier est aussi requis pour la saisie et la navigation. De nombreux retours visuels et sonores permettent de connaître les actions en cours. Ces retours sont assez importants car ils permettent d'avancer dans certaines «tâches». Dans 3D Webmaster, le retour visuel est bien utilisé, les mises en évidence à l'aide d'icônes et effets de surbrillance, offrent une meilleure compréhension des entités liées. Ils permettent ainsi de générer plus rapidement les actions voulues pour les entités. Les effets de mise en évidence permettent ici une visibilité de l'entité sélectionnée ainsi que de l'action en cours. Mais il n'est malheureusement pas possible de sélectionner correctement si un trop grand nombre d'entités sont présentes au même endroit. La sélection est une méthode et un support de communication. Divers protocoles existent qui simplifient les méthodes de sélection mais cette modalité de désignation reste néanmoins peu exploitée.

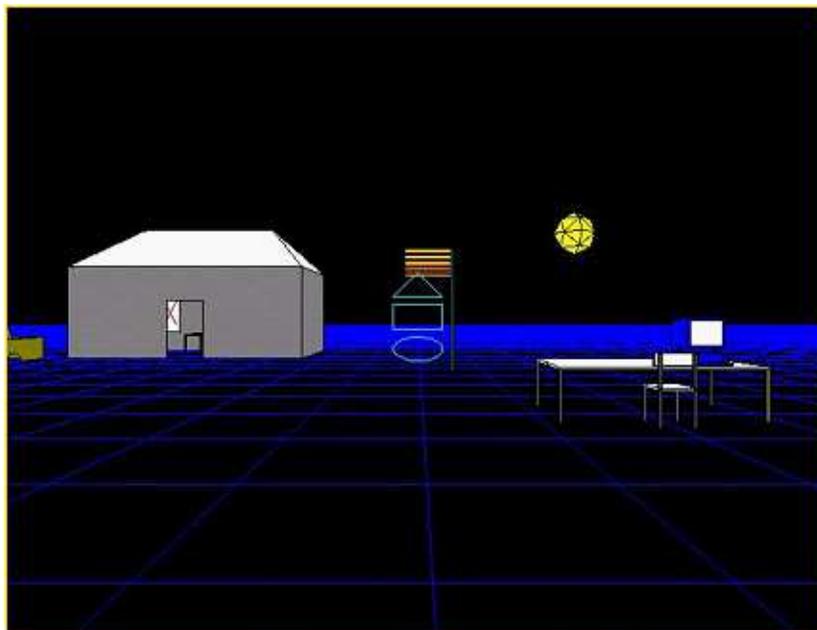
Conclusion

Nous avons vu que *SCL* introduisait tout comme dans *InViWo* [79] la notion de signal. Il permet une interaction non seulement des objets avec l'utilisateur mais aussi une interaction entre les objets. Cette forme de communication se rapproche de la réalité et contribue grandement à améliorer l'«ambiance» du *MV*.

3.4.5 Ulysse : agent conversationnel interactif

Les travaux de Pierre Nugues l'ont amené à la création d'un monde virtuel : *Ulysse* [73, 61, 7].

Ces travaux se sont portés sur la création d'un agent conversationnel permettant d'aider à la navigation dans les mondes virtuels. Il obéit aux requêtes de l'utilisateur en le déplaçant en conséquence. Pour ce faire, l'agent possède une base de connaissances sur le monde.

FIG. 3.12 – *Vue sur un monde d'Ulysse*

Analyse des types de communication

La communication entre l'utilisateur et l'avatar reste pour le moins simple. L'avatar est ici non-visible, il est néanmoins possible de le rendre visible par l'emploi d'extensions appropriées. L'agent reçoit les requêtes de l'utilisateur et agit sur le monde et l'avatar pour effectuer les actions voulues.

La relation entre le *MV* et l'humain peut se résumer par la création et la mise en place d'une table de symboles. L'utilisateur va définir des symboles pour les différents objets qu'il souhaite être connus par l'agent. Les symboles entrés, toutes les informations présentes au sein du *MV* vont être réutilisées pour permettre l'enrichissement de la base de connaissances de l'agent et ainsi augmenter l'éventail des requêtes reconnues.

La communication avec l'agent s'effectue dans les premières versions du programme à l'aide du clavier. L'utilisateur saisit les requêtes et l'agent répond par un «feedback» textuel et visuel avec les actions appropriées. Toutes les requêtes sont sous la forme de phrases, qui tendent à fournir de nouvelles positions et orientations à l'avatar, ce qui a pour conséquence de modifier la vision de l'utilisateur.

La communication est surtout orientée vers l'utilisateur d'une part à cause du langage utilisé et d'autre part à cause des intérêts visés par le projet.

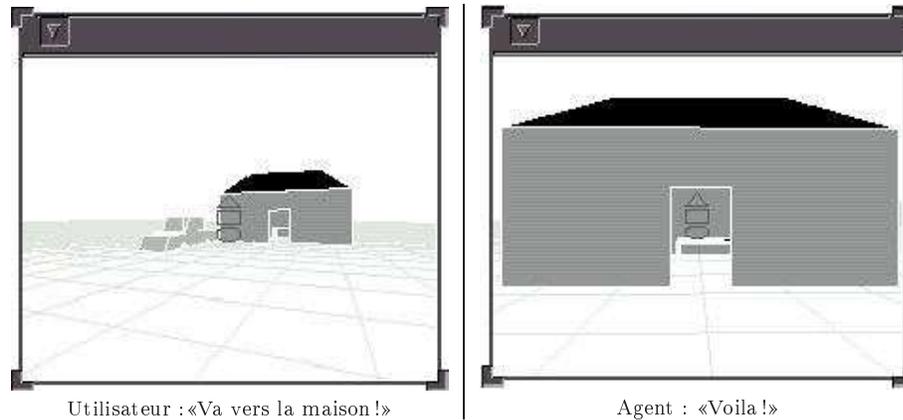


FIG. 3.13 – *Exemple d'une requête*

La communication entre les utilisateurs peut s'effectuer à l'aide de modules ajoutés au plug-in de visualisation mais aucun traitement n'est effectué lors des actes de communication entre les utilisateurs.

La communication avec le monde se limite surtout à l'entrée manuelle des symboles pertinents du monde virtuel. On effectue cette action par l'intermédiaire d'une fenêtre où nous devons rentrer toutes les informations concernant l'objet et ses points de vue. On se demande alors pourquoi nous ne pouvons pas utiliser l'agent pour rentrer ces informations pour des objets non connus. Ou ne peut-on trouver une méthode qui analyserait et généreraient ces informations réutilisables ? Il existe un élément de réponse : l'ensemble des variables du *MV* sont connues et demeurent accessibles. Dès lors, nous pouvons analyser le monde et en déduire les informations. Nous souhaitons réutiliser les informations présentes avant d'envisager une nouvelle table de symboles.

Conclusion

Les actions de déplacements s'effectuent à l'aide de différents supports. Les ordres énoncés au clavier ou encore dictés au microphone sont analysés pour être interprétés. Ainsi l'agent peut inférer sur l'action appropriée. On note que l'évolution du clavier au microphone a permis une saisie plus naturelle et rapide mais il présente une certaine imprécision qui peut une mauvaise action.

3.4.6 VREng : Virtual Reality Engine

VREng est une plate-forme de développement provenant des travaux des étudiants de l'ENST et dont Philippe Dax[99, 5] (<http://vrengr.enst.fr>) est à l'origine. L'objectif de cette plate-forme est prometteur, il s'agit de fournir un univers permettant le développement simple d'outils pour les mondes virtuels.



FIG. 3.14 – Vue du monde de VREng

Analyse des types de communication

A la date de décembre 2004 (version 6.3.0), il existe de nombreuses façons d'interagir avec le monde. Dans cet *EV*, les mondes virtuels sont regroupés suivant leur thématique.

La communication inter-utilisateurs ressemble fortement aux autres types de communication inter-utilisateurs. A l'aide de leurs claviers, les utilisateurs de *VREng* se transmettent des messages textuels pour dialoguer. Il n'existe pas d'autres possibilités de communication entre les utilisateurs.

Les sélections d'objets s'effectuent en pointant l'objet à l'aide de la souris. Après la sélection, une liste des actions possibles sur cet objet apparaît, nous pouvons ainsi modifier les éléments du monde et le faire réagir suivant nos

intentions. Les actions sont nommées et listées, nous pouvons grâce à ces noms connaître les actions définies sur l'objet sélectionné. La «sélection» est une des méthodes d'interaction entre l'utilisateur et le *MV*.

Un autre type de communication est plus interne à *VREng*. Elle traite de la communication inter-entités, d'un objet vers un autre objet ou même l'utilisateur. Nous pouvons affecter à un objet un code comportemental permettant de faire réagir l'objet en question à un événement donné. C'est une communication permettant de générer une certaine ambiance lors de la visualisation.

Conclusion

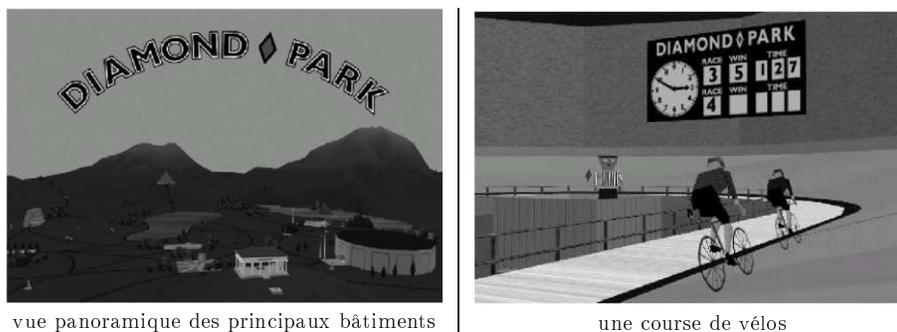
VREng présente une richesse d'interactions, source de collaborations et par ce fait, source de communication inter-utilisateurs. Il permet de nombreuses améliorations et une possibilité de structuration des données extensible. C'est une des raisons qui nous a poussé à choisir cet *EV* pour nos développements. Nous verrons cet *EV* plus en profondeur dans le Chapitre 5, page 122. La communication sous *VREng* entre les utilisateurs est très similaire à celle de *Dive*. Cependant, son intégration au monde est plus ergonomique.

3.4.7 Diamond Park : système social virtuel

Diamond Park est un environnement de simulation sociale. Il est important de signaler que cet *EV* n'a pu être testé par nos soins car son interface et son emploi nécessite un matériel informatique spécialisé. Toutes nos observations se basent sur l'étude des vidéos de démonstrations et des rapports techniques [105, 4]. C'est un environnement développé par «Mitsubishi Electric Research Laboratories» (*MERL*) pour la création d'un *EV* destiné aux rencontres et aux travaux collaboratifs.

Cette plate-forme a pour objectif d'offrir à des personnes géographiquement éloignées la possibilité de communiquer et/ou de participer à des activités communes. Nous pouvons y circuler librement et interagir avec d'autres visiteurs ou automates. Cet *EV* permet en outre de s'immerger dans un environnement 3D visuel et de manière assez innovante dans un environnement sonore en permanente évolution. Ce projet s'est orienté vers une tentative d'amélioration de la relation sociale virtuelle. Cette plate-forme invite les utilisateurs à communiquer et à partager des instants conviviaux.

Cet *EV* présente des formes de communication nouvelles et intéressantes. C'est ce que nous allons voir maintenant :



vue panoramique des principaux bâtiments

une course de vélos

FIG. 3.15 – Vues sur un monde de Diamond Park

Dans le futur, *MERL* souhaite construire différents environnements de simulation de vie sociale, permettant d'apprendre et de comprendre tout en se divertissant, soit un environnement de type *EIAH* d'apprentissage à l'aide d'*EVs*. De nombreuses études sont menées dans cette optique.

Ces travaux ont tout logiquement conduit à la création d'une plate-forme pour la construction de ce genre de systèmes appelée «Scalable Platform for Large Network Environments» (*SPLINE*). Ce système permet l'élaboration et la gestion de mondes sociaux.

Analyse des types de communication

Il existe trois types d'interactions à la bases des mondes gérés par ce système : Les interactions visuelles, les interactions physiques et enfin les interactions sonores. De toutes les interactions, les interactions visuelles et physiques sont similaires aux autres mondes de types *Dive*. Différentes activités existent et permettent une collaboration entre les différents utilisateurs du monde. Il existe aussi de nombreux automates permettant une amélioration de l'ambiance du monde. De toutes ces interactions, les interactions sonores restent les plus inhabituelles et innovantes.

Dès notre première entrée dans le Park de *Diamond*, nous sommes accueillis par les bruits des oiseaux et du vent, donnant cette impression champêtre de bienvenue. Dès lors, on associe instinctivement ce son à ce lieu d'entrée et de sortie. Les bruits et les sons nous fournissent constamment des informations sur le lieu où nous sommes et la fonction voulue par les créateurs. Le son est assez bien exploité pour communiquer certaines informations implicites. Il reste néanmoins des utilisations ambiguës du son qui donnent parfois une impression faussée.

Considérations artistiques

Les systèmes sociaux virtuels ont commencé à émerger comme une nouvelle forme d'art et plusieurs exemples existent déjà. Lorsque que le projet *Diamond Park* a été posé, il y avait deux lignes de développement visées : premièrement, laisser libre cours à l'imagination des utilisateurs et deuxièmement ; garder un environnement simple et facile d'emploi, pour une utilisation homogène de toutes les interactions possibles.

Diamond Park utilise une interface peu conventionnelle matérialisée une bicyclette. Ce choix de mettre tous les utilisateurs en bicyclette fût poussé par la volonté de faire interagir tout le corps et pas seulement une partie. L'autre raison plus philosophique consiste à améliorer l'interaction entre les visiteurs et par le biais du bicycle la rendre non-violente. Plusieurs démonstrations ont montré que la plupart des utilisateurs étaient charmés par cette utilisation et laissaient de côté leurs manières brusques. Ce support a permis une approche plus ludique du *MV*.

Ce choix a grandement influencé la création du monde avec de larges entrées et des rampes d'accès aux différents bâtiments, le *MV* pouvant ainsi se limiter aux bicycles seuls. Mais naturellement, par soucis de standardisation, les déplacements et quelques interactions considérés comme primitifs peuvent être effectués par sélection à la souris ou au clavier. Il est intéressant de noter que pour connaître l'interface employée par l'utilisateur, le *MV* représente différemment un utilisateur sur une interface standard – l'avatar apparaît sur un monocycle – et un utilisateur avec une interface spécialisée – l'avatar apparaît sur un bicycle. Nous voyons indirectement un souci de la part des programmeurs à faire connaître le contexte d'interactions de l'utilisateur. Il permet ainsi d'éviter les propositions inutiles d'interactions possibles seulement pour les utilisateurs de périphérie spécialisée. Cette visualisation du contexte interactionnel est un avantage pour les travaux collaboratifs où on comprend rapidement les possibilités des autres utilisateurs.

Le «design» de *Diamond Park* est intéressant, il se base sur une représentation idéologique entre la représentation graphique et sa fonction dans le monde. L'architecture des structures et leur position en fonction des autres permet une «pré-connaissance» des fonctionnalités et interactions disponibles par ces entités.

Conclusion

Le monde de *Diamond* présente une utilisation intéressante des *MV* avec une analyse constante du contexte de l'utilisateur. Dans un sens, *Diamond Park* tente d'améliorer la communication en prévenant des différents possi-



FIG. 3.16 – Vue sur un monde d'Active World

bilités du lieu présent. Il présente une application particulièrement innovante sur l'emploi des sons.

3.4.8 Active World : l'un des premiers EV

Active World (ou *Alpha World*) est l'un des projets de monde collaboratif virtuel les plus anciens[89, 81]. On estime le début de ce monde virtuel en ligne aux années 1990. Ce monde s'est rapidement développé pour atteindre aujourd'hui une taille comprenant près de 60 millions d'objets virtuels visualisables. Il est devenu un monde virtuel immense où les créations sont nombreuses et qui a su évoluer avec les airs du temps. On y retrouve des images et des posters datant des années 80. Pour la première fois dans cette étude, il nous a été possible d'observer réellement un monde «ancien» où les créations datent de plusieurs années.

Analyse des types de communication

De nombreuses interactions existent sous la forme de jeux. Malheureusement, la plupart ne sont disponibles que contre un paiement permettant d'obtenir le statut de «citoyen» du MV. La principale forme de communication utilise une fenêtre de «chat», où toutes les personnes peuvent s'échanger

des informations sans même tenir compte de la position des autres utilisateurs. Cela laisse encore une impression d'indépendance entre la fenêtre de chat et le *MV*.

Sinon les gestes et mouvements sont assez efficaces quoique limités. Ils permettent dans certaines conditions de rendre les actes de communication attrayants et clairs. De plus, les gestes et mouvements varient en fonction des fonctionnalités du monde. Dans un monde de batailles, les gestes disponibles sont différents d'attaques adaptées au *MV* de combat. De même, lorsque nous sommes dans le monde des «boîtes de nuit», toutes ces attaques se transforment en autant de pas de danse qui permettent aux plus habiles d'exécuter parfaitement un ensemble complexe de mouvements.

Mais dès lors, un phénomène inverse apparaît. Lors des dialogues inter-utilisateurs, certaines conversations sont tournées de manière à pouvoir utiliser ces mouvements. Ces gesticulations permettent ainsi de créer des situations amusantes. Étrangement, l'utilisateur cherche à utiliser au maximum ces mouvements rendant parfois les discussions difficile à suivre. . .

Les utilisateurs adaptent les discussions pour utiliser telle ou telle animation. Un utilisateur ayant connaissance des mouvements de joies(ex. mouvement de basculement du corps en arrière avec les mains sur le ventre) tournait toutes les conversations dans le but d'employer une maximum de mouvements.

La communication entre les utilisateurs s'effectue par une fenêtre de dialogue mais cette fois, dans cet *EV*, nous avons un retour visuel permettant de voir directement les messages sous la forme de bulles au-dessus de l'émetteur. Nous pouvons ainsi de manière plus simple faire le rapprochement entre le message et son émetteur. C'est une manière intéressante et innovante vis à vis des autres *EV*. Mais l'emploi de cette méthode est limitée à une communication proche avec peu de communicants. Lorsque nous sommes éloignés, les bulles sont illisibles et lorsqu'il y a de nombreux participants, il n'est plus possible de connaître son émetteur.

L'Environnement Virtuel d'Active World est orienté vers la découverte, la navigation et la création d'entités virtuelles. Il est intéressant de parler des créations. Prenons l'exemple d'un château situé dans un *MV* de constructions. Le «citoyen» propriétaire du château a souhaité créer un escalier en colimaçon pour accéder au sommet de son donjon. Pour ce faire, il n'avait à sa disposition que des objets divers et des cubes déformables. Il a pu utiliser ces cubes pour faire un semblant d'escaliers mais difficile néanmoins à utiliser. On remarque dès lors une similitude dans l'emploi des mouvements pour la communication et la création des constructions virtuelles. Le manque d'expressivités faciles conduit à des communications parfois ambiguës.

Conclusion

Nous terminerons avec cet *EV* en précisant que toutes ces créations proviennent pour la plupart d'une seule personne. Le mode d'interaction ne permet pas un travail collaboratif qui le soit vraiment. Les constructions s'effectuent de manière séquentielle. On peut le comprendre car les travaux sur les détections de collisions dans les *EV* distribués sont récentes. Cet *EV* permet une communication innovante avec les autres utilisateurs avec une utilisation de gestes et de mouvements contextuels.

3.4.9 Open Croquet : mondes virtuels collaboratifs

Open Croquet est une plate-forme sous licence LGPL ⁷ orientée vers la création d'applications multi-utilisateurs collaboratives en ligne. Une grande partie des recherches [94, 60, 59] sur cet *EV* s'est portée sur l'architecture qui permet de supporter une communication avec différents supports entre les utilisateurs, de partager des ressources ainsi que des calculs complexes. L'objectif de cette plate-forme est de permettre la création d'applications 2D ou 3D solides et à but collaboratifs.

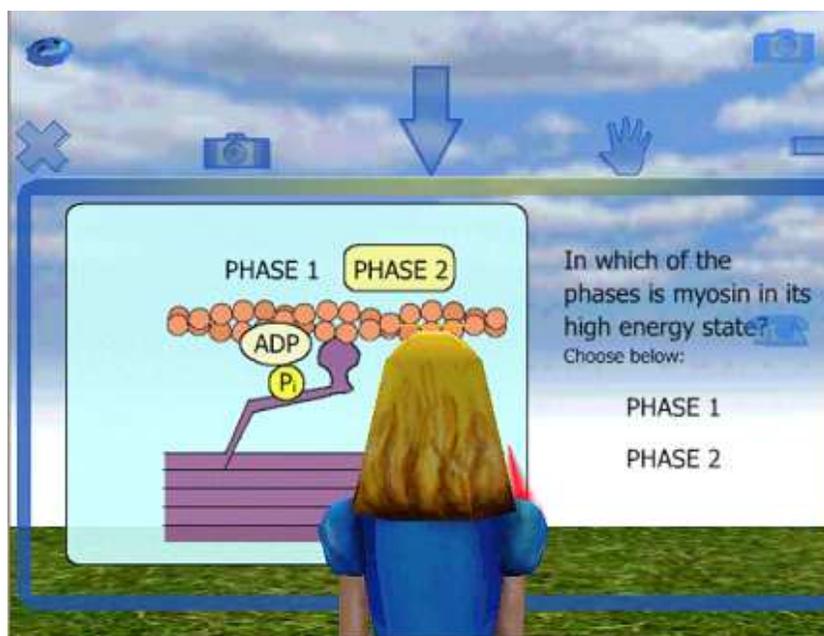


FIG. 3.17 – Vue sur un monde d'Open Croquet

Il est tout naturellement venu s'ajouter à notre étude à cause de son

⁷disponible à l'adresse : <http://www.opencroquet.org>

objectif clairement cité d'aider la collaboration entre les utilisateurs au sein d'un *MV* commun. Observons d'abord les communications possibles au sein de ce *MV* :

Analyse des types de communication

Notons en premier lieu que la communication entre les utilisateurs s'effectuent encore une fois sous la forme d'un «Chat» comme sous certains *EV* vus auparavant. Nous voyons cette fois une possibilité de dialoguer par la voix avec les autres utilisateurs. Mais là encore, avec cette méthode, il reste toujours le travail cognitif de faire correspondre la voix entendue et l'avatar dans la scène. Nous pouvons ajouter à cela que des utilisateurs débutants se retrouvent avec le même avatar permettant ainsi difficilement de faire correspondre l'avatar et la voix. Mais cela reste un problème redondant dans les *MV* actuels.

Nous passerons rapidement sur la communication entre l'humain et son avatar traitée qui est relativement bien soignée avec les orientations de la têtes et les diverses petites animations qui nous fournissent rapidement les actions en cours et nous permettent une certaine anticipation des actions voulues.

L'une des particularités intéressantes de ce monde est le partage de la désignation entre les utilisateurs. Elle est beaucoup plus efficace que dans *Dive* et permet une désignation avancée lorsqu'elle est couplée à la «représentation en miniature» du *MV*. On retrouve ici cette notion de «world in miniature» dans un des travaux de C. Chaillou[100]. En effet, dans *Open Croquet*, il est possible de générer un représentation réduite du *MV* simulée avec la possibilité d'utiliser cette représentation comme support à la désignation. L'utilisateur, lorsqu'il fait appel à cette méthode, peut observer devant lui à l'intérieur d'un cube, toute la scène visible ainsi que sa représentation dans la scène. Il peut dès lors tourner, déplacer et manipuler à sa guise le cube avec le monde à l'intérieur. L'utilisateur peut ainsi observer un autre utilisateur dans son contexte virtuel et partager dans une certaine mesure le même contexte graphique. Lors des actes de communication, il est alors possible de se référer au contexte de l'utilisateur pour mieux comprendre ses centres d'intérêts. La communication textuelle bien que peu innovante se retrouve devenir intéressante à l'aide de cette représentation réduite du contexte. Néanmoins, cette manipulation exige une certaine adresse à manipuler la souris pour orienter le cube comme nous le souhaitons.

Conclusion

La visualisation des actions et le «feedback» sont travaillés dans *Open Croquet*. Les travaux collaboratifs sont rendus plus aisés, notamment pour les manipulations où le contexte de la désignation est identifiable. Cet *EV* présente une base solide permettant de s'appuyer sur certains paradigmes intéressants.

3.4.10 Habbo hotel : un monde en ligne

Habbo est un *EV* simulant une communauté virtuelle qui favorise les rencontres entre amis ou d'autres utilisateurs de l'*EV*. Cet *EV* est un site Flash en ligne dont l'inscription gratuite est possible depuis 2004 (inscription et utilisation possible sur le site <http://www.habbo.fr>). Il a un objectif clairement annoncé de dialoguer librement avec d'autres utilisateurs et permettre des activités communes.



FIG. 3.18 – Vue sur un monde de Habbo

C'est un monde récent et particulièrement utilisé par des utilisateurs jeunes. Toutes les interactions sont faites pour être simples et faciles d'emploi. Observons en premier lieu les communications qui nous intéressent :

Analyse des types de communication

Cet *EV* est proche de celui d'*Active World* (cf. section 3.4.8 page 67) avec les mêmes objectifs mais cette fois avec une simplicité visuelle et interactionnelle. Néanmoins cette facilité existe au détriment d'une possibilité d'interactions avec le monde fortement diminuée. Les interactions possibles sont toujours (cf. *Active World*) adaptées au contexte de l'utilisateur. Ces interactions sont toujours liées au *MV* où l'avatar est présent.

Une des communications intéressantes de *Habbo* est la communication entre les utilisateurs. La méthode de dialogue aborde de nombreux aspects. Pour la première fois dans notre étude des mondes virtuels, cet *EV* simule la portée de textes envoyés. Dans les autres *EV*, un message était transmis à tout le monde. Toute personne du monde virtuel pouvait lire à tout moment l'ensemble des conversations. Dans certains, il était toujours possible d'envoyer des messages à des personnes en particulier sans que les autres n'aient connaissance. Mais cela était seulement une modification du destinataire/récepteur, aucuns traitements du ou sur le message n'étaient effectués.

Dans cet *EV*, il existe trois manières de dialoguer : «parler», «hurler» ou encore «murmurer». «Hurler» est une méthode similaire aux autres *EV* – le message est transmis à tous les utilisateurs du *MV*. Les deux autres méthodes font intervenir la notion de «portée». Lorsque nous dialoguons dans ces modes le message n'est transmis qu'aux utilisateurs présents à une certaine distance. Il simule d'une manière plus simple l'effet de la distance sur la voix – l'effet du bruit de Shannon sur le message (cf. chapitre 2 section 2.4.1 page 29). Les émissions de message ne sont plus visibles et compréhensibles au-delà d'une certaine distance. Néanmoins, nous recevons des bribes du message permettant de comprendre qu'un message a été murmuré.

Habbo présente ici une idée originale. Les créateurs sont allés plus loin en améliorant en plus la visualisation du message. Comme les messages émis manipulent la typographie du message pour donner des informations sur la nature de l'émission du message, les couleurs sont utilisées pour informer de l'origine du message (dans notre cas : l'auteur du message). De plus, la coloration du contour du message n'est pas laissée au hasard. Les vêtements de l'avatar sont analysés pour connaître la couleur globale perçue. Non seulement les habits de l'avatar sont analysés mais aussi la couleur des cheveux. D'après quelques observations, il apparaît que la couleur des cheveux a une plus grande influence que les autres éléments de l'avatar. Cette couleur perçue est réutilisée dans la coloration des messages et permet une désignation implicite de l'émetteur du message. Cette méthode possède elle aussi une certaine limite, notamment lorsque plusieurs avatars sont présents dans le *MV* et qu'ils tentent tous de dialoguer en même temps avec des couleurs proches.

Les développeurs ont voulu être au plus proche de la réalité en affectant un «temps de vie» aux différents messages transmis. Alors que les autres *EV* ont pour principe de sauvegarder toutes les communications. Cet *EV* ne permet pas aux utilisateurs de revenir dans les conversations, obligeant ainsi l'utilisateur à réitérer les actes de communications. Il permet ainsi d'éliminer les messages trop anciens automatiquement et fournit une information temporelle sur le message. Dans les autres *EV*, c'est une étape qui s'effectue manuellement dans les autres *EV* après la lecture fastidieuse de l'heure du message et l'heure actuelle. C'est là un choix qui comporte bien entendu des avantages de clarté et des défauts vis-à-vis du stockage des informations.

Conclusion

En conclusion, *Habbo* hotel est un *EV* intéressant où les créateurs utilisent de façon innovante des connaissances du *MV* pour l'amélioration de la communication entre les utilisateurs.

3.4.11 Spin-3D : pour un travail collaboratif

Le projet Spin 3D est un Environnement virtuel à but collaboratif (développé par France Télécom Recherche et Développement et le Laboratoire d'Informatique Fondamentale de Lille). « Cette plateforme permet à des utilisateurs distants de travailler ensemble, de produire des documents et d'interagir en temps réel dans un environnement virtuel. » (A. Galisson). Elle est destinée à assister un travail collaboratif. Elle est mise en œuvre dans différentes applications qui vont du domaine de la médecine (cf. figure 3.19) à l'apprentissage scolaire en passant par la modélisation de scènes virtuelles.

Analyse des types de communication

Les nombreuses applications de la plateforme ont un but collaboratif. De ce fait, l'*EV* est orienté vers les actes de communication entre utilisateurs. Dans cette plateforme, on retrouve une grande implication de l'*EV* dans les communications.

Dans l'application «Argonaute 3D» [55], des spécialistes de la médecine géographiquement éloignés peuvent dialoguer et discuter sur des cas pathologiques. L'intérêt de cette application est de partager un même point de vue sur un objet virtuel et de pouvoir ainsi dialoguer suivant ce contexte commun. Chaque manipulation de l'objet virtuel peut être transmise aux autres utilisateurs qui partagent ainsi un même point de vue. Ce contexte



FIG. 3.19 – *Vue sur Argonaute 3D*

commun a pour objectif de faciliter les prises de décision par une meilleure compréhension des messages échangées par les utilisateurs.

Suivant le même principe, nous avons les travaux sur le bureau virtuel 3D [100] où nous retrouvons l'idée de contexte virtuel commun. Dans cette application, la possibilité est offerte aux utilisateurs de visualiser le point de vue des autres utilisateurs suivant le paradigme de la boule de cristal. Un utilisateur peut ainsi transmettre précisément un centre d'intérêt à un collaborateur en proposant le partage de son point de vue.

Un autre type de communication entre utilisateurs apparaît dans l'application de téléconférence de Spin-3D : la communication non-verbale, plus précisément la communication gestuelle et la communication corporelle. L'objectif de cette application est de permettre à des utilisateurs de dialoguer autour d'une table virtuelle. L'*EV* assiste la téléconférence en fournissant un certain nombre d'informations sous la forme de communication non-verbale. Ainsi, lorsqu'un utilisateur rejoint ou quitte une discussion, l'*EV* se charge de communiquer cette information en effectuant une entrée ou une sortie de l'avatar de l'utilisateur de la table virtuelle. Ainsi, l'*EV* se charge de trans-

mettre aux autres utilisateurs des informations sous forme visuelle pour les autres utilisateurs. L'objectif est de faciliter la compréhension des différentes actions en cours.

Conclusion

On retrouve dans Spin-3D, un grand intérêt à fournir les informations d'un utilisateur aux autres utilisateurs. Il met en avant le besoin d'avoir un support commun pour effectuer un travail collaboratif. Dans un cas, les applications informent de certaines actions en cours ou et dans d'autres, elles offrent un support commun à la communication.

La transmission de certaines actions permet aux récepteurs de mieux comprendre l'environnement et les actions en cours. L'accès à un support commun permet aux interlocuteurs d'avoir un contexte commun pour communiquer. Les messages peuvent ainsi se fonder sur ce contexte pour transmettre les informations avec un risque mineur de mauvaise compréhension.

3.4.12 LEA : LIMSI Embodied Agent

La plateforme LEA, d'agents conversationnels développée par le Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur (LIMSI) et issue des travaux de J.C. Martin[44], permet l'élaboration d'agents. La particularité de cette plateforme par rapport à d'autres architecture d'agent conversationnel se situe au niveau de la compréhension de l'application par l'agent. En plus de pouvoir dialoguer avec l'utilisateur, l'agent apparaît alors comme mieux intégré à l'application (cf. figure 3.20). Il peut ainsi fournir des informations plus contextuelles.

Analyse des types de communication

Les communications se situent entre l'utilisateur et l'agent. La plateforme a pour but d'offrir des agents virtuels capables d'assister l'utilisateur dans l'emploi d'une application. Le premier principe utilisé est celui du «Chat-terbot» : l'utilisateur écrit et envoie une requête textuelle à l'application. L'agent l'intercepte, tente de la comprendre afin de pouvoir y répondre. Ainsi lorsque l'utilisateur émet un souhait ou une question, l'agent essaye de comprendre et de répondre à son attente. J.P. Sansonnet [57] a trouvé intéressant que l'agent puisse répondre à ces requêtes de manière isolée en tenant compte de la seule requête, puis idéalement de répondre en suivant

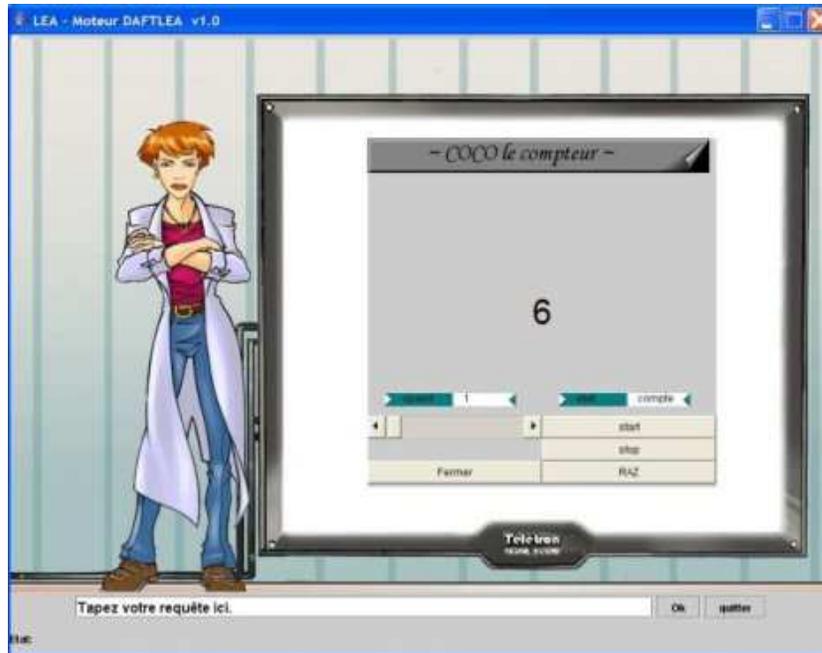


FIG. 3.20 – Vue sur le système DaftLEA

toute une fenêtre de dialogue.

Dans la plateforme LEA, lorsque l'utilisateur émet une requête, par exemple lorsqu'il souhaite connaître l'emplacement d'un élément sur la fenêtre d'application, l'agent fournit l'information de manière textuelle ou orale mais aussi s'il est visible de manière gestuelle. On voit apparaître la notion de redondance des informations. Une communication qui serait à la base pauvre : envoi d'un texte, est enrichie par l'emploi du support visuel (les gestes). Ces informations visuelles viennent parfois à compléter une information textuelle. La compréhension du message s'en trouve améliorée. C'est la juxtaposition des informations qui va fournir l'information souhaitée par l'agent.

Conclusion

De toute évidence, il apparaît que l'emploi de plusieurs canaux de sorties améliore les possibilités de compréhension du message dans le sens souhaité. Suivant G. Gerbner, la compréhension d'un message dépend de la manière dont on y accède, or on remarque qu'il est difficile de prévoir l'emploi de tel ou tel canal d'accès pour obtenir la communication souhaitée. Les travaux de

S. Buisine [15] montre cette hétérogénéité entre les personnes dans l'emploi et la visualisation des informations. Dès lors, nous revenons à la question (page 36) sur les possibilités du choix du support en fonction du contenu pour une meilleure compréhension du message.

3.4.13 Divers mondes virtuels

Il existe de nombreux autres mondes et environnements virtuels. Cette panoplie d'*EV* reste bien réduite face à la création quasi continue de nouveaux *MV*. Nous présentons brièvement quelques *MV* intéressants que nous avons laissé de côté pour éviter certaines redites. Voici donc une autre liste non-exhaustive d'*EV*.

Second Life

Il nous était impossible de ne pas citer ce monde virtuel, source de d'études pour de nombreuses communautés. Plusieurs millions de personnes à travers le monde se connectent sur ce monde virtuel afin de vivre et de se fabriquer une «seconde vie». Les objectifs de ce monde sont orientés vers la création, c'est tout naturellement que le monde regorge de possibilités de créations. Des bureaux virtuels d'IBM ou encore de partis politique peuvent dorénavant être visités sur la place principale du monde.

La grande force de cet *EV* réside dans la possibilité quasi infinie d'interactions possibles, chaque mouvement, chaque action peut être paramétrable. Un langage, le «Linden Script Language» (*LSL*)⁸ est apparu pour permettre non seulement aux experts mais aussi aux néophytes de développer une partie personnelle de leurs interactions. De plus, ce langage permet de créer de véritables jeux dans le monde virtuel. Un wiki intéressant a vu le jour et recèle une mine d'informations pour connaître Seconde Life. -> <http://secondlife.com/badgeo/wakka.php?wakka=HomePage>.

Sims Online

Les Sims Online est une version du célèbre jeu «Les Sims». Dans «Les Sims», le but du jeu était de gérer un personnage (un «Sim»), ou plus généralement une famille de Sims. Le rôle du joueur consiste à guider le ou les Sims tout au long des journées. Nous devons lui faire faire les tâches usuelles de la vie courante – ex. le faire manger, dormir, se laver, se divertir... Les possibilités du jeu sont assez importantes, ce jeu est doté de nombreuses extensions qui augmentent les possibilités d'interactions. Un des aspects du jeu

⁸Il est proche sous certains aspects de *SCL* mais présente une richesse importante.



FIG. 3.21 – *Vue sur un monde de Second Life*

réside dans la possibilité de nouer des relations entre les Sims s'échelonnant de la haine viscérale à l'amour passionné.

Mais évidemment, nouer des relations avec un ordinateur ou avec ses propres personnages limite quelque peu le jeu. C'est pourquoi une version du jeu Les Sims avec laquelle il est possible de jouer avec d'autres joueurs réels est apparu : «Les Sims Online».

Du point de vue interactionnel, on retrouve les interactions existantes vues précédemment. La richesse de ce *MV* réside dans les possibilités simples mais diversifiées d'interactions entre les utilisateurs. Une utilisation intéressante qui se retrouve dans les autres *MV* (notamment *Habbo hotel*) est l'utilisation des objets virtuels pour faire varier les interactions disponibles.

On peut ajouter à cela la manière de communiquer les émotions sous la forme d'icônes représentatifs rendant simples leur emplois. Cette méthode ajoute à la communication par «Chat», une richesse de communication.

Le village 3D

Le village 3D a été créé en 1998. Ce village virtuel est géré par les habitants du *MV*. Il permet de nous incarner en un personnage qui peut se déplacer, danser, dialoguer. . . Nous avons aussi la possibilité d'interagir ges-



FIG. 3.22 – Vue sur un monde des sims online

tuellement avec les autres villageois, par exemple en saluant de la main. Ce monde est presque une copie de Active World mais avec des interactions et des méthodes de dialogues légèrement différentes s'approchant de *Habbo hotel* avec une grande amélioration apportée par la 3D.

Les interactions sont similaires. Nous pouvons nous promener dans le Village 3D et découvrir les centaines de maisons construites par les villageois ainsi que les bâtiments collectifs : la mairie, la plage, la piscine... Nous pouvons ainsi rendre visite aux villageois et visiter leurs maisons, défiant parfois certains lois de la gravité que la physique ne pourrait tolérer.

En outre, il est possible de partager des médias comme écouter de la musique dans la discothèque ou encore sur les chaînes Hi-Fi des villageois ou les JukeBoxes d'un Bar de la ville. Il existe surtout une multitude de jeux auxquels il est possible de participer : des courses de voitures, des quizzes permettant de recevoir des lots, une chasse aux trésors et bien d'autres choses.

VR4all

VR4All nous permet de visiter des univers 3D en évolution constante, d'admirer les constructions issues de compositions d'artistes et des internautes eux-mêmes.

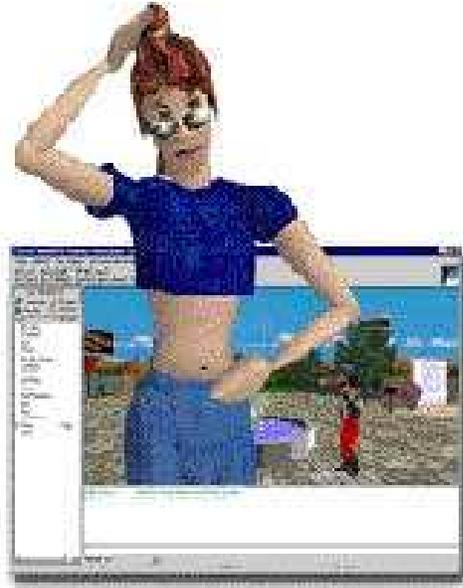


FIG. 3.23 – Vue sur un monde de *Village 3D*

Comme dans les autres *EV*, nous pouvons discuter, faire connaissance mais VR4All permet aussi de partager des images, des vidéos ou des musiques avec les autres internautes de la communauté.

Les communautés de VR4All se construisent petit à petit. Tout comme dans Alpha World, nous pouvons investir dans la vie d'une communauté existante ou construire sa propre communauté. Cet *EV* présente certains aspects d'interactions intéressantes comme les dialogues de groupes mais n'apporte pas de réels nouveautés.

Les possibilités de VR4All sont nombreuses, nous citerons les plus utilisées :

- Création de jeux en ligne, de RPG⁹, MMORPG¹⁰, jeux d'aventure ;
- Architecture, création rapide de maquettes, visite de sites virtuels, reconstitution de sites ayant existés ;
- Magasins en ligne, galerie d'artistes virtuelles ;
- Education – compréhension de la 3D, jeux de compositions ;
- Reconstitution de scènes pour la criminologie ou le théâtre.

⁹RPG = diminutif de «Role Player Game», jeux de rôle

¹⁰MMORPG = diminutif pour «Massive Multiplayer Online Role Player Game», jeux de rôle en ligne massivement multijoueur



FIG. 3.24 – Vue sur un monde de VR4all

Les MMOG

Les MMOG, les «Massive Multiplayer Online Game» (les jeux online massivement multijoueurs) sont une source d'études intéressante. On peut donner une liste de plus d'une centaine de programmes et logiciels permettant une communication médiatisée par un monde virtuel. Voici pour terminer cette étude, des vues sur quelques MMOG où les communications présentent des aspects innovants mais quelques peu éloignés de notre axe de recherche qu'est la communication inter-utilisateurs.

3.4.14 Synthèse de l'étude

Les types de communication

Notre étude révèle qu'il existe de nombreux types de communication. Certains environnements virtuels sont centrés sur les actes de communication entre l'utilisateur et son Avatar (cf. Ulysse section 3.4.5 page 60). Il existe d'autres *EV* qui s'orientent vers une communication entre les Objets(cf. le langage SCL 3.4.4 page 58 ou encore InViWo[79]). De tous ces types de communication, nous nous sommes intéressés aux mondes virtuels traitant de la communication entre les interlocuteurs.



Bzflag



Kochonland



Ryzom



SpaceCowboy

FIG. 3.25 – Divers MMOG

Le paradigme du Chat

Un premier constat a été de remarquer l'utilisation massive par les mondes virtuels de l'outil de «Chat» (cf. caractéristiques du «Chat» dans l'analyse du *MV Dive* 3.4.3 page 55). La communication entre les interlocuteurs s'effectue au moyen de messages textuels.

Diversification du support

Certains *EV* ont cherché à enrichir cette communication en donnant différents supports, notamment les mondes à but collaboratif : *Dive* ou *Croquet* (cf. section 3.4.9 page 69). Ils permettent la visualisation de centres d'intérêts à l'aide de «représentation en miniature» et de traits de désignation. D'autres comme *Active World*, *Habbo*, *Second Life* ou *VR4all* permettent une communication à l'aide de mouvements ou de gestes contextuels au *MV* (cf. *Habbo* section 3.4.10. page 71).

Enrichissement du texte

Plusieurs *EV* effectuent des traitements sur le texte lui-même. *Active World* (cf. section 3.4.8 page 67) affiche par exemple directement le message sur la scène virtuelle et permet ainsi une meilleure compréhension du contexte. Il améliore l'utilisation du «Chat». *Habbo* (un monde très orienté vers la communication interpersonnelle) fini d'intégrer complètement le texte à la scène en utilisant différents traitements sur le texte. Il effectue des traitements sur l'émetteur pour améliorer la compréhension de la provenance du texte (la coloration des textes est liée aux vêtements de l'émetteur). Il propose différents modes de dialogues : parler, hurler ou murmurer, qui simulent la portée de la voix. *Habbo* utilise fortement les informations virtuelles pour améliorer des composantes de la communication (cf. *Lasswell* Chapitre 2.4.2 page 32).

Recherche d'expressivité

L'analyse des mondes virtuels et leurs évolution montre une recherche de leurs auteurs à améliorer la communication dans les mondes virtuels. Ils offrent de plus en plus de moyens de fournir et d'obtenir des informations pour la compréhension d'un acte de communication.

3.4.15 La RV et la communication

De cette étude, nous sommes à même de voir les possibilités offertes par la réalité virtuelle pour intervenir sur les actes de communication. Dans cette partie, nous allons tenter de répondre à chacune des questions de communication afin de résumer les possibilités nouvelles et spécifiques qui permettent d'améliorer la communication dans les mondes virtuels.

Question 1 :

Avec quelle exactitude les symboles de la communication peuvent-ils être transmis ? (cf. section 1 page 30)

Réponse :

Contrairement au monde réel qui subit les aléas du «bruit» perturbant la perception du monde, la *RV* offre une certaine solidité sur ce point de vue. En effet, la solidité des réseaux assure une protection contre ce phénomène de «bruit», tout «symbole» exprimé (cf. Shannon et Weaver) dans le monde virtuel est transmis sans altérations vers le récepteur. La réalité virtuelle permet une assurance sur la solidité de la transmission.

Question 2 :

Avec quelle précision les symboles transmis véhiculent-ils la signification désirée ? (cf. section 2 page 30)

Réponse :

Cette question pose le problème l'expressivité du sens désiré par le biais d'outils virtuels. C'est une question difficile car la finesse des symboles à transmettre et le sens désiré sont difficilement quantifiables.

Il n'existe pas de systèmes universels en réalité virtuelle. Certains environnements virtuels s'orientent vers un objectif précis. Ils permettent de transmettre parfaitement certaines significations dans un contexte restreint. La réalité virtuelle peut aider par des systèmes d'interfaces homme-machine spécialisées à transmettre une information sous un format adapté.

Question 3 :

Avec quelle efficacité la signification reçue affecte-t-elle la conduite dans le sens désiré ? (cf. section 3 page 30)

Réponse :

Cette question revient à se poser la question de la correspondance entre le sens désiré et le sens reçu. Lorsque l'on regarde le modèle de Shannon, on observe page 31 que le modèle de communication adapté aux *MV* montre une difficulté et même une incapacité du *MV* à pouvoir agir et influencer sur le sens désiré. Le Monde virtuel reste une interface entre l'émetteur et le récepteur : il peut influencer sur les questions techniques ou fournir des apports quant aux problèmes sémantiques de la communication. Malheureusement, les questions sur l'efficacité restent dans le cadre de la *RV* difficilement formalisables.

Question 4 :

Quelles sont les composantes pouvant être améliorées pour obtenir au final une amélioration de la communication ? (cf. section 4 page 33)

Réponse :

Le modèle de H. Lasswell page 32 propose un découpage intéressant de la communication. Observons les améliorations possibles des composantes par le biais d'un *EV*. Nous remarquons que les environnements virtuels traitent de chacune d'entre elles.

des différents composantes. Chacune des composantes traite d'un domaine différent et toutes peuvent subir un traitement en vue d'améliorer la communication.

Question 5 :

Le contexte est une source d'informations supplémentaires mais quand et comment peut-on l'utiliser afin qu'il soit un apport à la communication ? (cf. section 5 page 36)

Réponse :

G. Gerbner (cf. Chapitre 2.4.3 page 34 nous apprend que les événements du monde réel ne peuvent être perçus dans leur intégralité. Seule une partie des informations qui proviennent de l'événement peut être perçue et analysée par le récepteur.

L'emploi des mondes virtuels dans les actes de communication montre une différence importante à ce niveau. En effet, le monde simulé possède la faculté de fournir, d'une façon ou d'une autre, tous les paramètres des différents événements en cours. Un événement généré par le monde virtuel peut aussi être analysé dans son ensemble. Dès lors, il est possible de fournir tous ces paramètres aux entités de l'*EV*.

Question 6 :

Quelle forme choisir pour un message connu ? (cf. section 6 page 36)

Réponse :

Nous avons vu dans ce chapitre les limites de la *RV* (cf. page 51), la préférence d'un type d'information plutôt qu'un autre est un choix difficile. En effet, il est difficile de choisir une forme « parfaite » pour un contenu informel. De plus, le contexte diffère souvent d'un individu à un autre. Par exemple, le modèle des Rileys (cf. Chapitre 2.4.6 page 39) sur les groupes d'influence et de contexte entre les utilisateurs nous montre que chaque forme choisie dépend du contexte du récepteur.

La *RV* est un outil qui peut offrir différents moyens de visualiser ces informations. Elle nous fournit plusieurs formes visuelles et différentes formes textuelles possibles que nous verrons par la suite. Il apparaît alors intéressant d'offrir la possibilité de laisser l'utilisateur choisir la forme la plus adaptée au contenu. Une approche consiste à offrir au récepteur la possibilité de choisir le support voulu pour comprendre le message transmis.

Question 7 :

Le contexte de la communication ne permet t-il pas de prédire certains actes de communication ? (cf. section 7 page 37)

Réponse :

Le contexte de l'utilisateur apporte des informations. Les modifications de contexte peuvent donner une idée sur l'ensemble des actions possibles et donc sur les actions pouvant suivre. Nous ne pouvons prédire les actions futures mais nous pouvons en proposer suivant le contexte. Durant les actes de communication, un émetteur énonce des informations vers un récepteur dans l'attente d'une action. Nous pouvons tenter de comprendre le message et proposer d'effectuer l'action attendue.

Question 8 :

Quand et comment les «agents-clés» peuvent-ils aider la communication à être plus pertinente ? (cf. section 8 page 39)

Réponse :

Nous pouvons modéliser ces «agents-clés» comme étant une partie intégrée au monde virtuel. De ce fait, ils ont la possibilité d'accéder aux informations du monde virtuel. Du fait que les messages transitent par le monde virtuel, ces «agents-clés» peuvent accéder aux actes de communication.

T.H. Newcomb décrit la communication comme étant plus pertinente lorsqu'elle permet aux communicants d'atteindre plus rapidement leur situation commune (cf. Newcomb chapitre 2.4.4 page 37). Pour rendre la communication plus pertinente, il faut aider la communication à atteindre cette situation.

Question 9 :

Comment se définissent les groupes d'influences qui nous permettent de comprendre les communications ? (cf. section 9 page 40)

Réponse :

Les groupes d'influences peuvent se modéliser en réalité virtuelle comme différentes sources de références pour les locuteurs. La notion de groupe d'influences raffine les idées de G. Gerbner sur les informations contextuelles. Cette idée de groupes se retrouvent en réalité virtuelle à de nombreux niveaux(ex : contexte visuelle ou contexte interactionnel). Ces contextes ou sources d'informations, sont souvent invisibles et non connus pour les communicants. Nous pouvons fournir ces contextes et ces informations au récepteur.

Ces informations permettent de mieux comprendre les actes de communication reçus.

3.5 Conclusion

Si sophistiquée soit-elle, la machine ne saurait acquérir les compétences dialogiques de l'Homme et, par exemple, comme le rappelait Descartes, répondre adéquatement à toute question dans n'importe quel contexte. Par contre, il est parfaitement concevable d'envisager de modéliser une interaction avec une machine dédiée à une tâche déterminée. (Préface D. Vernant)- Interaction et pragmatique [J. Caelen].

De l'analyse des mondes virtuels, il ressort que la communication est une composante forte de la réalité virtuelle. Certes, les types de communications sont variés et chaque système de réalité virtuelle possède ses spécificités. Les orientations et les objectifs de chaque système tendent à le spécifier et à le particulariser mais les environnements virtuels ont des caractéristiques communes dont la connaissance des informations du monde simulé. De l'analyse de la réalité virtuelle confrontée aux problèmes de communication, il ressort que le support virtuel représente un support de communication avec une source d'informations commune. L'analyse de cette source et des actes de communication peut permettre une meilleure compréhension des messages par les interlocuteurs.

C'est ce que nous allons expliquer dans le prochain chapitre de la communication en définissant la communication améliorée et sa mise en œuvre.

4 La communication améliorée

Sommaire

4.1 Le concept d'amélioration	90
4.1.1 Définitions	90
4.1.2 Nos champs d'actions	91
4.1.3 Exemples d'améliorations	96
4.1.4 Conclusion	99
4.2 Mise en œuvre des améliorations	100
4.2.1 L'agent	100
4.2.2 Cas d'utilisation	104
4.2.3 Les modes opératoires de l'agent	111
4.2.4 Réalisation des améliorations	114
4.2.5 Conclusion	119
4.3 Conclusion	120

La chose la plus importante en communication, c'est d'entendre ce qui n'est pas dit. [Peter Drucker]

Le chapitre 2 relatif à la communication a soulevé de nombreuses interrogations. Certaines questions de communication sont restés avec des réponses partielles, nous pouvons mentionner les problèmes cités Shannon-Weaver (cf Section 2.4.1 page 29). La réalité virtuelle se présente à nous comme l'outil de communication le plus évolué. Nous avons vu à travers notre étude (cf. Section 3.4 page 53) qu'elle permet de résoudre localement des problèmes de communication grâce à la connaissance du monde virtuel.

Dans un premier temps, nous allons présenter une définition de ce que nous entendons par communication améliorée. Nous spécifierons notre champ d'action et présenterons les exemples d'améliorations que nous souhaitons obtenir. Dans un second temps, nous présenterons la mise en œuvre de nos améliorations qui a conduit à l'élaboration d'un agent conversationnel. Nous définirons son rôle et son action sur la communication.

4.1 Le concept d'amélioration

4.1.1 Définitions

Le concept de communication améliorée est nouveau. Pour expliquer notre approche, nous allons rappeler le point de vue de H. Lasswell (cf. Section 2.4.2 page 32) qui présente la communication comme une subdivision en six éléments :

- L'intention : à l'origine du message ;
- L'émetteur ;
- Le message : l'information à transmettre ;
- Le médium : le support de la communication ;
- Le récepteur ;
- L'impact : l'effet obtenu après la réception du message.

La communication améliorée est la résultante d'une amélioration d'une ou de plusieurs de ces composantes.

Dans cet ensemble de composantes, nous avons choisi de nous concentrer sur : **le message** et **le médium**.

Le message

A word to a wise is sufficient.[Dicton anglais]

Les messages et les informations transmises peuvent être incomplets ou ambigus. Dans nos usages de dialogues, de vis-à-vis ou médiatisés, il est commun d'omettre certains détails surtout s'ils apparaissent inutiles ou triviaux. Les conséquences de ce choix sont parfois importantes et conduisent à une incompréhension¹ ou une compréhension faussée du message. L'acte naturel humain pour résoudre ce problème est de raffiner l'acte de communication par un ajout d'informations au message initial lorsqu'il est possible, ou par une levée des ambiguïtés existantes.

Le médium

Language is the medium of thought.[Hegel]

¹cf. Problème des groupes d'influences décrit par Riley [82], cf. Chapitre Communication Section 2.4.6 39

Le médium d'un acte de communication est vu comme une interface qui peut être définie par un groupe de processus et destinée à restituer un message. Les interfaces «formattent» les intentions de l'émetteur avant de les transmettre. Symétriquement, la réception d'un message s'effectue par des interfaces de réception qui «perçoivent» les informations pour être traitées par le récepteur. On observe ainsi l'un des problèmes majeurs cité par Shannon-Weaver et G. Gerbner (cf. Chapitre 2 pages 29–34) concernant la disparité entre les symboles transmis et les symboles reçus par les interfaces. Sans détailler ces problèmes, on comprend que l'amélioration de la communication pourrait résulter d'une amélioration des supports de transmissions au niveau de l'émetteur (l'expressivité, Shannon) ou au niveau du récepteur (accès à l'informations, Gerbner).

4.1.2 Nos champs d'actions

Comme nous l'avons dit plus haut, notre recherche d'amélioration de la communication porte sur une amélioration de deux composantes de la communication : le message et le médium. Nous allons essayer de définir nos champs d'actions en étudiant ces composantes et en comprenant les éléments qui les composent.

Le message

Notre première étude se centre sur l'analyse du contenu et plus précisément l'analyse du contenu textuel des messages transmis dans les environnements virtuels. Nous souhaitons comprendre la composition des messages textuels.

En règle général, lorsque les phrases sont correctement formulées, la taille des phrases est un critère de lisibilité. Elle indique la richesse syntaxique et sémantique du texte.

Nous n'aborderons pas ici les problèmes linguistiques. Nous nous limiterons à justifier les choix techniques. Nous noterons que les formules de lisibilité (cf. Flesh et Gunning[12]) fondées sur le rapport entre la compréhension et la longueur des phrases prédisent que le nombre moyen de mots par phrase intervient dans la difficulté de lecture. Un texte est plus facile à lire si ses phrases sont courtes. La longueur ne devrait pas dépasser 15 à 20 mots.

Mais les dialogues entre personnes ne respectent pas ces règles. Des études ont été menés [95, 27, 14, 13]. En effet, le travail d'écriture d'un article ou d'un livre (d'une thèse...) diffère des actes de communication en temps réel qui ont leurs spécificités. Une observation a été menée sur un serveur de

discussion où les personnes dialoguent et s'échangent des messages suivant le protocole de communication « Internet Relay Chat » (*IRC*). Le serveur est contrôlé par des super-utilisateurs qui en limitent les abus. Nous avons observé quatre types de salons de discussion :

1. Le 1^{er} est un salon de discussion où un ChatterBot a été placé. Le ChatterBot permet de tenir une conversation plus ou moins cohérente avec un ou plusieurs interlocuteurs suivant un système d'apprentissage (notation sur le graphique : Bot).
2. Le 2^e est généraliste et regroupe des conversations de type « Chat » que l'on retrouve habituellement sur ces serveurs (notation sur le graphique : Chat).
3. Le 3^e salon se situe dans un monde virtuel où les participants doivent se coordonner et manipuler des objets virtuels ensemble pour atteindre un objectif (notation sur le graphique : MV collaboratif).
4. Le 4^e salon est un monde peu collaboratif où les dialogues ne sont pas de première importance pour l'accomplissement des objectifs du *MV* (notation sur le graphique : MV peu collaboratif).

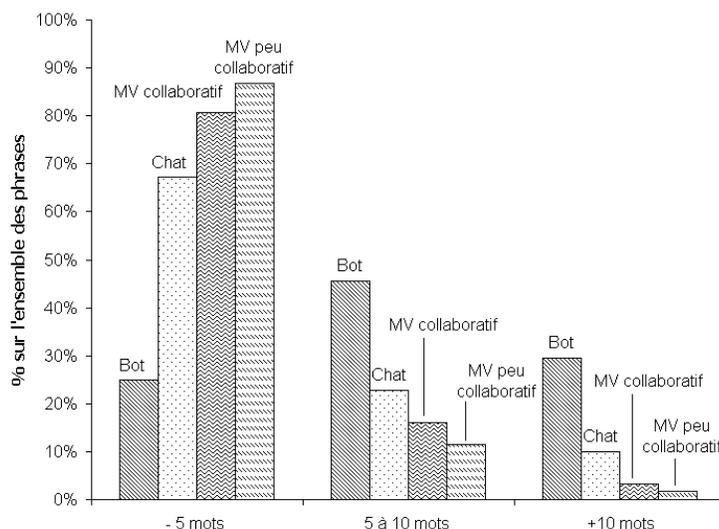
L'étude a été d'estimer la longueur des phrases utilisées par les utilisateurs. Le corpus utilisé représente cinq échantillons de 100 000 phrases prises au hasard dans la base des messages stockés respectivement sur les serveurs : *irc.u-psud.fr* et *gameboy.fleming.u-psud.fr*². Pour chaque salon, nous avons analysé automatiquement à l'aide de scripts PERL les messages transmis entre les utilisateurs. L'analyse comporte un traitement sur le nombre de mots utilisés par les usagers. Après cette analyse, Les résultats³ obtenus sont présentés Figure 4.1.

longueur	Bot	Chat	MV 1	MV 2
< 5 mots	25%	67%	81%	87%
5 à 10 mots	40%	23%	16%	11%
> 10 mots	29%	10%	3%	2%

TAB. 4.1 – Tableau des statistiques

²Tous les droits d'accès à ces messages appartiennent respectivement à l'association Aurore et Flemnet. Les droits accordés ont été seulement de pouvoir accéder à ces messages et de lancer les scripts sans droits de citations des messages. Protection par les droits de vie privée.

³présentation des résultats autorisée par les administrateurs Aurore et Flemnet.

FIG. 4.1 – *Histogramme sur la longueur des phrases*

Analyse

Nous remarquons que les messages longs (plus de 15-20 mots) sont inefficaces (cf. Tremblay [98]). Les messages de plus de 15 mots sont presque inexistantes. Lorsque l'utilisateur a conscience de la présence du bot, il effectue un effort pour être compris du bot et utilise des phrases longues pour que celui-ci puisse comprendre. Sur le tableau, on voit que près de 30% des messages dépassent les 10 mots. Mais lorsque les utilisateurs dialoguent sans assistant virtuel, ils emploient des phrases souvent courtes quel que soit le contexte. Sur les résultats, 90% des messages de type «Chat» général ont une longueur inférieure à 10 mots. Dans les *MV*, la proportion des messages inférieurs à 5 mots est importante, elle atteint 81% pour les mondes collaboratifs et 87% pour les mondes non-collaboratifs.

On en déduit que les discussions sont limitées et se résument à quelques mots dans les mondes collaboratifs et peu collaboratifs. C'est un fait auquel nous nous attendions pour le monde peu collaboratif mais moins pour le monde collaboratif. Les utilisateurs de ce monde collaboratif utilisent aussi des phrases courtes. Lorsque l'on analyse en détails les jeux de dialogues, on observe bien des phrases courtes mais elles ont la particularité d'être très pertinentes. Pour se coordonner et agir, les utilisateurs emploient un système propre de mots qui, placés dans le contexte du monde virtuel, ne présentent pas d'ambiguïtés. Lorsque l'on étudie les messages en détails, on note qu'ils sont composés en majeure partie de sujets, verbes, objets ou encore de petites informations complémentaires.

Conclusion

Cette analyse nous informe sur un point important : les utilisateurs de mondes virtuels emploient peu de phrases longues (de plus de 10 mots). Nous pouvons alors conclure que la majeure partie des actes de communication est composée des phrases courtes (moins de 5 mots). Dès lors, le message peut être décomposé en éléments simples avec une limite pour la décomposition.

La polysémie

On peut lire dans les propos de J.F. Dortier[24] concernant la communication qu'«elle ne pourra jamais être totalement objective et sans ambiguïté». Il restera toujours dans un acte de communication une part d'ambiguïté. Nous aborderons dans cette partie une des difficultés soulevées par le contenu des messages : la polysémie.

Une des limites réside dans la difficulté à formuler des messages clairs et explicites. La polysémie des termes contribue à rendre ambiguës certaines communications. L'usage de termes à sens multiples et significations implicites fait partie de notre « jeu » de communication usuel. Cette pluralité des significations est un des obstacles à l'amélioration de la communication. Selon, J.F. Dortier, elle est mise en évidence par Charles S. Peirce, Roland Barthes, Umberto Eco. Dans le cas d'une phrase à caractère polysémique, quel sens devons-nous choisir ? Existe-il des dominances permettant de déterminer le sens ? Existe-il des cas spécifiques qui tendraient à réduire notre champ d'action ? Ces questions restent en suspens.

L'incommunicabilité ? C'est pas qu'on ne communique pas assez. On communique trop et mal.[Robert Lalonde] - La belle épouvante

Le médium

Notre seconde étude s'est portée sur le contenant. Lors de l'analyse des communications des mondes virtuels (cf. Section 3.4.1 page 55), nous avons étudié les différents cas d'actes de communication avec les différents supports utilisés. Pour les communications inter-utilisateurs, nous observons que nous observé que la majorité des mondes virtuels emploient le support textuel et dans une moindre mesure le support déictique. Nous allons étudier ces supports afin de connaître le médium de la communication et en définir leurs caractéristiques.

Le support textuel

Le support textuel possède ses spécificités, ses avantages et ses limitations. Les systèmes Ulysse[7] et WordsEye[23] peuvent utiliser l'interface au clavier pour saisir les informations sur les objets. Dans le premier système, nous pouvons discuter avec un agent et obtenir des indications sur un objet ou son point de vue. Nous pouvons alors accéder à des informations sans que l'objet ne soit visible. Dans le second système, nous pouvons agencer les objets les uns en fonction des autres avec de simples mots. C'est un des avantages du support textuel : accéder et manipuler rapidement un objet. Mais nous voyons apparaître des limitations : le premier ne permet pas d'accéder à des zones inconnues (dont la dénomination est inconnue) et le second ne permet pas le placement à des positions imprécises (dont le placement n'est pas dictable) (cf. explication plus détaillée par J. Caelen [17]). De même, lorsque les caractéristiques distinctives entre les objets ne permettent pas une différenciation pertinente, il est impossible avec un emploi du clavier de désigner un objet. On peut citer le cas de la sélection d'un galet sur une plage de galets. Dans une telle situation, l'emploi de la modalité textuelle requiert des « mots » pour désigner un galet précis. C'est une des faiblesses du support textuel[14].

Le support déictique

Un second support dit déictique, pour la désignation, est souvent employé dans les mondes virtuels. Il permet de pointer une zone d'intérêt ou un objet visible. Ce support déictique permet une interaction intuitive (l'interaction à l'aide de la souris a été étudiée par Strommen[97]). On remarque dans cette étude qu'elle permet une adaptation rapide. C'est l'une des raisons de sa présence dans la plupart des interfaces des *MV*. L'utilisation de la souris est importante si nous souhaitons une manipulation directe et rapide. Elle constitue un avantage par rapport au support textuel qui requiert plus de temps. Par contre, elle possède aussi ses limites. Une de ces limitations, citée par Fraser[34], est liée au champ de vision. Le support déictique nécessite une vision de la zone d'intérêt. Pour aller au-delà de cette limite, le support textuel apparaît comme complémentaire.

Vers un système bimodal

De ces propos découle l'idée d'employer ces deux modalités (cf. définition d'une modalité section 4.2.2 page 104) de façon complémentaire. Les avantages de l'une nous apparaissent compenser les limitations de l'autre. La communication dans les *MV* utilise ces modalités mais il reste des commu-

nications seulement textuelles ou seulement déictiques. Cette pluralité des modalités améliore l'expressivité de l'émetteur. La communication qui restait difficile dans certaines conditions s'en trouve facilitée. Ainsi la mise en place d'une fusion bimodale (ou multimodale de manière plus générique) élargit nos champs d'actions.

L'ajout du support déictique dans les actes de communication est une amélioration déjà utilisée dans de nombreux travaux. On peut distinguer dans ce contexte plusieurs types d'utilisations multimodales dans les environnements virtuels. Les uns utilisent une fusion multimodale pour déplacer, manipuler, observer les objets[22, 96, 27]. Les autres utilisent ces modalités de manière indépendante – cf. Croquet 3D[94]. Enfin, d'autres encore considèrent le support déictique comme simple outil d'annotation ou de désignation – cf. VREng[5] ou *DIVE*[36].

Dans le cadre de la communication inter-utilisateurs, nous avons la possibilité d'élargir l'expressivité de l'émetteur durant les actes de communications. Ce fait constitue à nos yeux une amélioration du support et donc de la communication.

Ecrire est une forme de mensonge. C'est-à-dire de fiction, de hâblerie, de mystification. Simplement parce qu'on ne peut pas tout dire. [Robert Lalonde] La belle épouvante

4.1.3 Exemples d'améliorations

Après avoir défini nos champs d'action au niveau du contenu et du médium, nous présentons notre objectif global qui est d'assister les communications entre utilisateurs. Nous voulons intervenir au moment des actes de communication. Par améliorations, nous entendons trois actions :

1. étayer les informations disponibles ;
2. prévenir les ambiguïtés ;
3. répondre aux attentes.

Nous allons expliquer nos souhaits en matière d'amélioration dans chacun de ces cas.

Étayer les informations

Lorsque les utilisateurs communiquent entre eux, le message présenté sous une forme apporte des informations à l'utilisateur. Le but de la communication étant une mise en commun des informations, l'émetteur va chercher à fournir des informations pour y parvenir.

Si nous reprenons les idées de la communication suivant le point de vue de T.M. Newcomb, nous avons retenu que la communication avait pour objectif d'atteindre une situation commune (cf. Chapitre 2.4.4 page 37). C'est le rôle des informations contenues dans les actes de communication. Fournir plus d'informations permet une convergence plus rapide vers une situation commune. La communication devient alors plus efficace d'où une communication améliorée. Si nous reprenons le scénario, nous souhaitons ajouter à certains messages plus d'informations.

Exemple : « Je préfère Chez-Pierre. ». Nous voulons pouvoir fournir des informations sur cet objet : « Chez-Pierre ».

Comment y parvenir ?

Notre souhait est d'augmenter les informations disponibles dans le message. Pour y parvenir, nous avons choisi d'analyser l'objet concerné par le message et de fournir différentes informations sur celui-ci.

Qu'est-ce qui caractérise un objet ?

Un objet possède de nombreuses caractéristiques : une forme, des couleurs, une position, une orientation et même une histoire. Dans le réel, lorsque l'on accède à un objet, on accède à une partie de ses caractéristiques. Dans le virtuel, il est possible d'accéder à toutes ces informations.

Au delà de ses propres caractéristiques, un objet peut aussi être caractérisé par sa relation aux autres entités virtuelles. Ainsi les informations relatives aux autres objets est une autre source d'informations que nous pouvons exploiter.

Nous essayerons alors de comprendre l'objet de l'acte de communication (« Chez-Pierre » dans notre exemple) et fournir des informations sur celui-ci. Ces informations se situent aussi bien au niveau du contenu (informations textuelles) qu'au niveau du médium (informations déictiques/visuelles).

Prévenir les ambiguïtés

Lors des actes de communications, il n'est pas rare d'omettre des détails. Ces omissions peuvent aboutir à des conséquences graves qui peuvent éloigner la communication de son point de convergence. Il est possible dans certains cas de prévenir cet éloignement. Une seconde possibilité d'amélioration concerne les moyens d'éviter les ambiguïtés possibles.

Les connaissances sont contextuelles à un utilisateur ou un groupe. L'émetteur génère son message dans ce contexte avant de le transmettre. M.

and J. White Riley et G. Gerbner ont précisé l'existence de tel contexte (cf. Chapitre 2.4.6 et 2.4.3). Par ailleurs, ces auteurs précisent que l'existence d'un contexte commun permet une meilleure compréhension. Dans notre cas, nous avons ce contexte commun : l'environnement virtuel. Il nous reste à savoir l'utiliser correctement pour arriver à nos fins.

Par exemple, dans le scénario d'introduction au manuscrit, nous voyons A indiquer à E : « Je t'attends au café. ». On devrait être en mesure de comprendre que ce message cache une ambiguïté et intervenir pour prévenir l'émetteur des possibles incompréhensions.

Comment y parvenir ?

Nous souhaitons aider la communication en révélant les ambiguïtés. Lorsque les actes de communication concernent un objet de l'environnement virtuel, nous voulons comprendre le message et analyser le *MV* pour savoir si le message ne contient pas des ambiguïtés. Nous souhaitons analyser les informations géographiques et visuelles du monde virtuel par rapport à l'émetteur et au récepteur pour évaluer l'ambiguïté du message.

Répondre aux attentes

La dernière amélioration que nous souhaitons apporter consiste à répondre aux actions de l'émetteur. Toujours en suivant les théories de T.M. Newcomb, nous souhaitons aider la convergence vers une situation commune (cf. page 96), en fournissant directement une réponse aux attentes de l'émetteur.

Certains messages de l'émetteur peuvent se présenter suivant deux cas :

- des questions
- des ordres

Ces «attentes» requièrent une intervention, soit informelle en réponse à une question, soit interactionnelle par une manipulation. Nous souhaitons intervenir pour répondre à ces attentes.

Toujours en reprenant le scénario initial, nous voyons E poser une question : «Où est le café Chez-Pierre ?». Nous comprenons que l'émetteur souhaite connaître sa position géographique et nous en connaissons la réponse. Nous sommes apte à répondre à cette attente.

Comment y parvenir ?

Les messages peuvent contenir une question ou une requête. Nous souhaitons analyser le message pour connaître son contenu sémantique. Nous voulons vérifier les attentes de l'émetteur et lui offrir la possibilité d'obtenir la réponse à ses attentes.

4.1.4 Conclusion

Les améliorations que nous souhaitons effectuer sont orientées vers le traitement des informations. Les environnements virtuels nous offrent la possibilité d'effectuer ces traitements. Ils possèdent en toutes les informations utiles. Il nous faut trouver un moyen de les extraire et de les exploiter. Ces traitements et ces souhaits sont à l'origine du concept de «communication améliorée».

4.2 Mise en œuvre des améliorations

Après avoir expliqué la notion de communication améliorée et nos souhaits, nous allons nous appliquer à trouver des méthodes pour la mise en œuvre de ces améliorations.

Cette mise en place s'est effectuée en plusieurs étapes. La première a été l'élaboration d'un agent conversationnel pour analyser et comprendre les messages. Nous avons par la suite étudié les cas possibles d'interactions entre les locuteurs. De ces cas d'utilisation, nous avons défini les interventions possibles de l'agent. Enfin, Nous présenterons les améliorations retenues.

4.2.1 L'agent

Les besoins de traitement des messages nous ont conduit à l'élaboration d'un agent conversationnel. Ce dernier a pour rôle de récupérer et d'analyser les messages et d'agir en conséquence. Pour comprendre son rôle et son champ d'action, nous allons présenter la notion d'agent. Nous pourrions par la suite analyser la portée de notre agent et définir les actions de celui-ci en fonction du message.

La notion d'agent

La définition de Jacques Ferber[29] est celle qui a retenu notre attention et qui semble être la plus à même de répondre à nos exigences. J. Ferber caractérise un agent comme une entité physique ou virtuelle :

1. qui a la capacité d'agir dans un environnement ;
2. qui a la capacité de communiquer directement avec d'autres agents ;
3. qui est mue par un ensemble de tendances ;
4. qui possède des ressources propres ;
5. qui a la capacité de percevoir son environnement et qui ne dispose que d'une représentation partielle de cet environnement ;
6. qui possède des compétences ;
7. qui offre des services ;
8. qui peut éventuellement se reproduire.

Cette définition diffère de notre conception d'un agent dont voici nos caractéristiques suivant ce point de vue :

1. Nous souhaitons qu'il agisse sur l'environnement notamment lorsque l'émetteur est dans l'attente d'une action.

2. Nous n'avons pas traité notre approche dans un contexte multi-agents.
3. Nous l'avons mue d'un ensemble de rôle où l'agent tend à offrir des possibilités d'interactions.
4. Nous avons souhaité qu'il possède des propres ressources notamment concernant le traitement linguistique.
5. Nous avons souhaité utiliser toute la perception possible de l'agent de son environnement sans jamais le limiter dans sa vision.
6. Notre agent possède des compétences de traitements.
7. Il offre la possibilité d'étayer les informations disponibles.
8. Nous n'avons pas recherché cette possibilité.

Agent de recherche d'informations

Par le terme d'agent de recherche d'informations, nous qualifierons toute la gamme de logiciels intermédiaires entre les moteurs de recherche et les «agents intelligents», consacrés à la recherche d'informations. Leur intérêt réside dans leur capacité à remplir et à automatiser des tâches. Ils suivent, à la lettre, la définition du terme agent : «entité agissant pour le compte d'un bénéficiaire» [85, 87].

Cette définition, quelque peu vague au premier abord, est admise dans de nombreuses recherches [85, 87]. Les agents de recherche d'informations sont d'une grande diversité et ne remplissent pas, pour la plupart, les mêmes tâches. Nous pouvons, malgré cela, les distinguer par les fonctionnalités qu'ils possèdent. Gardons en tête que ces distinctions tendent à s'estomper devant les avancées technologiques qui intègrent des fonctionnalités de plus en plus complexes dans un même ensemble.

On distingue quatre fonctions principales qui caractérisent les agents intelligents :

1. La recherche d'informations : celle-ci peut se faire de manière «intelligente» par l'utilisation de moteurs préexistants et perfectionnés (ex : Xpath, XQuery pour la recherche dans des fichiers XML), d'outils d'analyse linguistique des requêtes (ex : Lex et Yacc pour la reconnaissance lexicale et syntaxique) ou par inférences des informations engendrées à partir d'autres informations accessibles (cf. moteur d'inférences, analyse géographique) ;
2. L'analyse des informations récupérées : indexation des résultats, résumé automatique ;
3. Le filtrage, l'édition, l'archivage, la mise à jour de résultats ;
4. La visualisation des résultats.

Les technologies utilisées sont variées, intégrant parfois des technologies issues de l'intelligence artificielle. Ces fonctions sont à la base de notre agent mais là encore les possibilités sont nombreuses. Dans le domaine des agents de recherche, ces agents intelligents ont attiré notre attention et nous avons décidé d'en approfondir les spécificités.

Les agents intelligents

Qu'est-ce qu'un agent intelligent ? : Objet utilisant les outils et techniques de l'intelligence artificielle : il adapte son comportement à son environnement en mémorisant ses expériences et se comporte comme un sous-système capable d'apprentissage : il enrichit le système qui l'utilise en ajoutant, au cours du temps, des fonctions automatiques de traitement, de contrôle, de mémorisation ou de transfert d'information. AFNOR – Association Française de NORmalisation.

Ainsi si nous rassemblons les informations, un agent est considéré comme intelligent s'il contient un ou plusieurs des éléments suivants :

- Une base de connaissances pré-définies ;
- Un moteur d'inférences, fournissant un support à des raisonnements complexes ;
- Un système d'acquisition de connaissances ;
- Un mécanisme d'apprentissage.

Caractéristiques des agents intelligents

On s'accorde sur la nécessité de l'existence de quelques caractéristiques pour qu'on puisse parler d'agents intelligents. A l'heure actuelle, il n'existe pas d'agent qui rassemble toutes ces caractéristiques. Voici un aperçu des attributs définis par les chercheurs Oren Etzioni, Daniel Weld, Stan Franklin et Art Graesser[33] :

- **L'autonomie** : l'agent doit pouvoir prendre des initiatives et agir sans intervention. Dans le contexte de notre recherche, il doit pouvoir agir alors que l'émetteur n'exprime pas le désir de communiquer. C'est le cas par exemple d'Autonomy[75].
- **Capacité à communiquer et à coopérer** : l'agent doit pouvoir échanger des informations plus ou moins complexes avec d'autres agents, avec d'autres programmes ou avec des Humains. Le service Firefly ou Intellishoper[28] permet, par exemple, à l'agent représentant un locuteur, d'échanger des informations avec d'autres agents représentant d'autres locuteurs afin de rechercher ceux d'entre eux ayant les mêmes

centres d'intérêts, dans le but de les rapprocher et de les faire communiquer.

- **Capacité à raisonner, à réagir à leur environnement** : l'agent doit être capable de s'adapter à son environnement (qui peut être composé d'autres agents ou de programmes externes) ainsi qu'à ses évolutions. Cette adaptation doit obligatoirement s'appuyer sur une analyse de l'environnement extérieur.

Caractéristiques de notre agent

De toute évidence, notre agent ne regroupe pas toutes ces caractéristiques. Nous avons souhaité à l'aide de cette définition des agents intelligents définir les caractéristiques de notre agent :

- **L'autonomie** : nombreux éditeurs de logiciels «agents» contournent le problème de l'autonomie par modifications de leurs agents afin qu'ils puissent, à intervalles réguliers, effectuer les tâches confiées par l'émetteur. C'est là un cas où on ne peut véritablement parler d'autonomie car l'agent n'est pas en permanence « à l'écoute » et ne fait que reproduire, comme un simple automate, les demandes préprogrammées par le locuteur. Notre agent utilise aussi ce contournement. Il n'est pas actif à tout instant et ne s'active que ponctuellement (lors d'un acte de communication). L'agent intercepte la communication pour effectuer une série de traitements permettant d'obtenir la communication souhaitée.
- **Capacité à communiquer et à coopérer** : Nous pouvons citer les travaux de N. Richard[80] où les activités se sont concentrées sur cette communication inter-agents. Chaque agent coopère et permet un accroissement des connaissances. Il apparaissait important que notre agent communique avec l'émetteur. A titre d'exemple, les actes de communication initiés par l'émetteur peuvent comporter une ambiguïté. Si les connaissances du monde et de l'émetteur peuvent lever cette ambiguïté ou une partie de cette ambiguïté, il est utile d'effectuer les traitements adéquats. Une communication s'effectue avec les différentes parties du moteur de simulation et si besoin avec l'émetteur pour améliorer les informations de l'acte de communication.
- **Capacité à raisonner, à réagir à leur environnement** : l'agent doit être capable de s'adapter à son environnement (qui peut être composé d'autres agents, de programmes externes) ainsi qu'aux évolutions de celui-ci. Cette adaptation doit obligatoirement s'appuyer sur une

analyse de l'environnement extérieur.

Conclusion

Notre agent est un agent de type conversationnel, il analyse et il comprend les actes de communications. Il possède une base d'information qui regroupe l'ensemble des informations du monde virtuel. En outre, il a la possibilité de converser avec l'émetteur d'où son nom : Agent Conversationnel Intelligent (ACI). Pour comprendre ses modes opératoires, nous allons étudier différents cas d'utilisation pour couvrir et expliquer l'ensemble des actions possibles.

4.2.2 Cas d'utilisation

L'élaboration de notre agent faite, nous nous sommes penchés sur l'analyse des différents cas d'utilisation. L'émetteur utilise ses modalités pour communiquer. Dans les environnements virtuels étudiés, les modalités sont employées de façon indépendante. Pour notre système, nous avons souhaité qu'il existe une synergie dans l'emploi de ces modalités. Cet emploi synergique des modalités fait partie intégrante du domaine de la multimodalité. La fusion multimodale concerne le traitement en entrée des informations et la fission concerne le traitement en sortie des informations.

Définition d'une modalité

Notre recherche s'est ainsi dirigée vers l'étude de deux médias : un dispositif de saisie textuelle et un dispositif de désignation graphique. Ces médias sont les dispositifs physiques utilisés pour l'acquisition des données nécessaires aux modalités. L. Nigay [71] propose une définition de la notion de media qui peut être synthétisée en : dispositif servant de support à l'information.

En quoi consiste une modalité? Les définitions proprement académiques se retrouvent dans les écrits de L. Nigay[71]. L. Nigay assimile la modalité à une technique d'interaction. Pour notre part, nous utiliserons la définition de J.C. Martin[64]. Il définit la modalité de manière générale comme un processus d'analyse de données. Quoique succincte, cette définition nous permet d'introduire le schéma fonctionnel d'une modalité (cf. figure 4.2).

Dans notre cas, une modalité est définie par des données D analysées à un instant donné, un processus (P) (ou programme), un ensemble de résultats (R), un ensemble de conditions (C_e) permettant un contrôle des données en

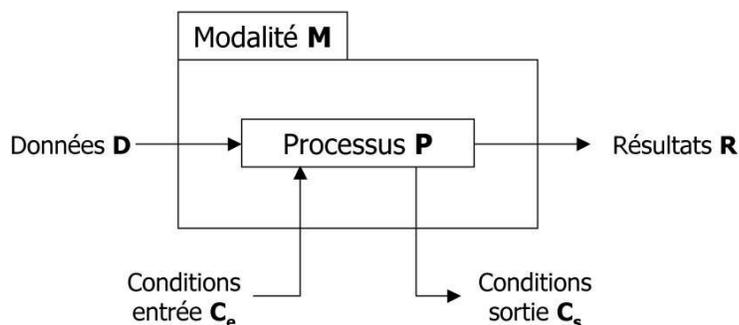


FIG. 4.2 – Schéma fonctionnel d'une modalité

entrée et un ensemble de conditions (C_s) permettant d'effectuer un filtre des informations en sortie.

Fusion multimodale

Dans cette partie, nous étudions la multimodalité (cf.[22, 95, 65, 47, 92, 106]) relative à deux modalités. De ce fait, le terme multimodalité peut sembler trop large pour notre recherche. Par la suite, nous emploierons indifféremment le terme «multimodale» ou «bimodal».

En premier lieu, nous nous sommes concentrés sur les possibilités d'employer ces modalités en analysant leurs champs d'actions. Par la suite, nous avons étudié les cas d'utilisation de la fusion des informations (fusion multimodale). Enfin, nous nous sommes canalisés sur la manière d'afficher et d'utiliser les modalités pour fournir des informations lisibles et compréhensibles.

Les champs d'actions des modalités

Nous avons d'abord tenté de définir les champs d'actions du couple (textuelle-déictique). Nous avons vu que leurs champs d'actions se complètent idéalement. Que se passe-t-il au niveau des liens entre ces modalités ? C'est ce que nous allons essayer de comprendre. L'étude des modalités s'est effectuée dans le contexte de notre environnement de développement VREng (cf. Chapitre 5.1 page 122).

La modalité textuelle

La modalité textuelle que nous associons dans notre cas aux entrées par l'interface clavier, possède des caractéristiques pauvres et peu développées. L'observation de sa décomposition nous donne :

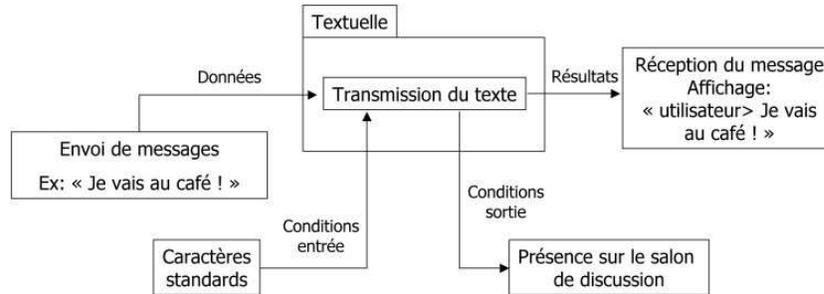


FIG. 4.3 – *La modalité textuelle*

Cette décomposition (cf. figure 4.3) nous indique plusieurs points à noter. Seules quelques analyses basiques (Qui est l'émetteur ? Quel est le canal de discussion ?) sont effectuées par le processus de la modalité. Le traitement du message est inexistant. Lorsque l'on compare les données en entrée et les résultats, aucun traitement notable est effectué.

Nous voulons mettre en place un agent, ce qui laisse supposer que le processus de traitement soit modifié. Nous allons supposer que notre agent est en place et que les traitements sont opérationnels. Notre décomposition est alors modifiée par l'agent et nous obtenons la figure 4.4.

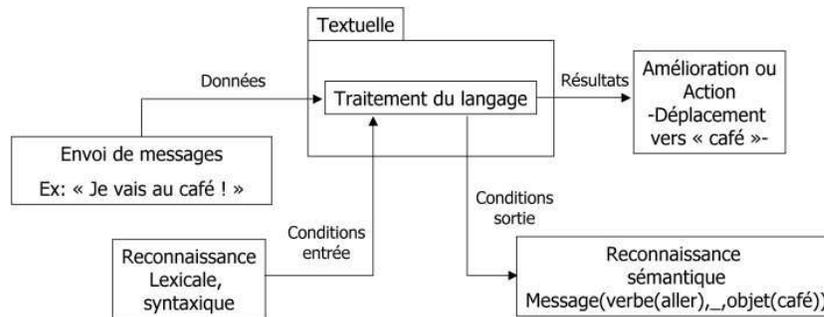


FIG. 4.4 – *La modalité textuelle améliorée*

La modalité textuelle que nous souhaitons obtenir améliore le traitement des données. L'agent effectue un traitement linguistique (cf. figure 4.4). Le texte entré est analysé à l'aide d'un moteur linguistique capable de reconnaître une partie des informations du message. Cet apport de l'agent a

modifié les propriétés de cette modalité. Il nous est maintenant possible de comprendre le fonctionnement possible de cette modalité avec notre seconde modalité. Nous allons effectuer la même analyse pour la modalité déictique.

La modalité déictique

La modalité déictique est souvent utilisée pour désigner gestuellement un objet ou une personne. Son emploi se cantonne à ce seul objectif (cf. Communication dans les *MV* Chapitre 3.4.1 page 55). Regardons ses composantes au sein de notre système (cf. figure 4.5).

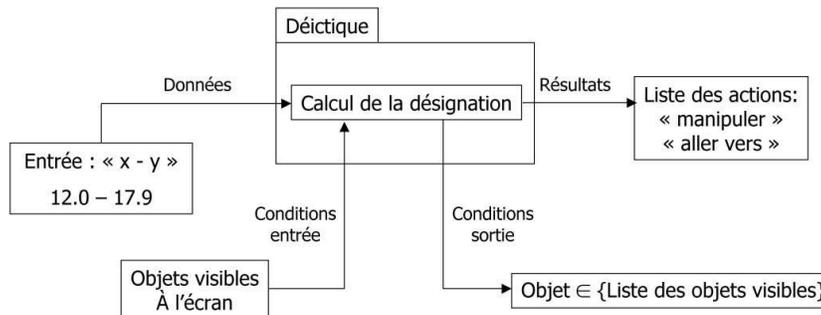


FIG. 4.5 – *La modalité déictique*

Une observation nous indique que l'entrée des données se résume à des coordonnées (x, y) . On remarque alors la faible portée de nos souhaits dans la mise en place de l'agent. Nous souhaitons aider l'émetteur dans son emploi de la désignation. Si nous supposons notre agent en place et nos améliorations faites, nous obtenons la décomposition de la figure 4.6.

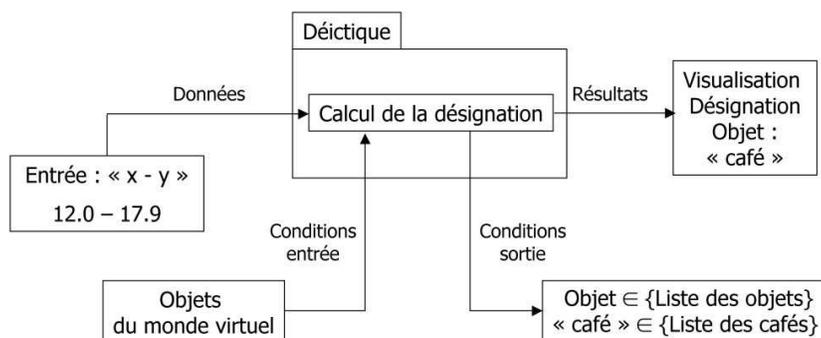


FIG. 4.6 – *La modalité déictique améliorée*

L'expressivité de l'émetteur avec la modalité déictique reste faible. Mais

nous avons facilité la désignation avec des méthodes de désignation et de visualisations. Nous obtenons alors des conditions d'entrée et de sortie plus larges. La mise en place de notre agent et de nos méthodes a permis d'améliorer les possibilités d'actions.

Une modalité principale

Nos constats se rapprochent des remarques de J.Caelen et A. Xuereb sur :«Interaction et Pragmatique»[17] (à la section 6.1.2 page 227–235) concernant les usages de la multimodalité. Pour la fusion des informations, nous avons observé les différentes informations transmises par les différentes modalités. Si nous observons les modalités, nous remarquons que la modalité textuelle est plus «expressive» que la modalité déictique. La modalité textuelle offre la possibilité d'exprimer plus d'informations. Les objectifs exprimés par la modalité textuelle sont plus diversifiés. Même si nous avons amélioré la modalité déictique, elle reste limitée et ne permet pas la transmission d'informations complexes comme la modalité textuelle.

Au niveau des interactions avec ces modalités, la modalité textuelle ne peut être analysée qu'au moment de l'envoi du texte. L'émetteur risque peu de se tromper sur ce qu'il souhaite envoyer avec la modalité textuelle. Si l'émetteur écrit un texte incorrect, il a la possibilité de le réécrire. Ce n'est pas le cas de la modalité déictique dont le traitement s'effectue au clic de la souris et où l'émetteur peut difficilement revenir sur cette désignation. A partir de ce constat, nous avons étudié différentes méthodes[46, 31, 96, 10] et nous avons choisi de considérer la modalité textuelle comme principale lors des actes de communication.

Emploi multimodal

Nous présentons maintenant les différents cas d'utilisation multimodale :

1. Modalité textuelle seule : « Va vers le bar! »
2. Utilisation déictique seule : **désignation d'un centre d'intérêt**
3. Redondante : « Va à droite! » **désignation à droite**
4. Utilisation complémentaire : « Va vers ce café! » **désignation du café**
5. Contradictoire : « Va vers ce café! » **désignation de l'épicerie**

Dans les deux premiers cas (1,2), les informations ne proviennent que d'une seule modalité. C'est ce qui existe actuellement dans de nombreux environnements virtuels.

Le cas de redondance (cas **3**) est difficilement applicable à notre agent. En effet, les traitements de l'agent se déclenchent à l'envoi du message. Or, si nous sommes arrivés à comprendre le message, l'emploi de la seconde modalité est alors considéré comme un second message et non un complément du précédent. Il est difficile dans notre contexte de traiter ce cas. Lorsque la modalité textuelle possède toute les informations nécessaires, le traitement s'effectue sans tenir compte de la modalité déictique.

L'utilisation complémentaire (cas **4**) est le cas que nous souhaitons obtenir. Une utilisation « combinée » permet un emploi synergique des modalités. L'utilisation indépendante des modalités nous limite dans certains actes de communication. Nous souhaitons utiliser les informations déictiques lorsque les informations textuelles sont insuffisantes pour la compréhension. Ce cas complémentaire soulève le besoin d'une seconde analyse au niveau de l'arrivée des informations pour différencier les cas d'utilisation.

Le cas contradictoire (cas **5**) est une source d'ambiguïtés. L'agent analyse le message et ne comprend pas le lien entre l'objet textuel et l'objet désigné. L'agent fournit un réel apport, il révèle cette contradiction. Nous avons essayé d'aller plus loin et d'empêcher l'apparition de ce cas. L'existence de ce cas prouve que l'émetteur après l'envoi du message textuel peut désigner un objet non concerné par le message. L'agent va prévenir l'apparition de ce cas en limitant le champ de sélection. L'émetteur ne peut plus sélectionner tous les objets mais seulement ceux que l'agent considère comme possibles.

Ces cinq cas regroupent l'ensemble des cas de fusion des informations pour un système bimodal. Par notre choix de considérer une modalité principale, nous avons pu simplifier et choisir les actions appropriées par l'agent. Le cas complémentaire (cas **4**) soulève un problème que nous n'avons pas étudié : l'ordre d'utilisation des modalités.

Cas d'utilisation temporels

Dans le cas d'une utilisation complémentaire, nous pouvons obtenir d'un point de vue temporel. Trois cas se présentent :

1. **Simultané** : « Va vers ce **désignation** café ! »
2. **Alterné pré-fixé** : **désignation** « Va vers ce café ! »
3. **Alterné post-fixé** : « Va vers ce café ! » **désignation**

Le cas n°1 consiste à désigner à la souris pendant l'écriture au clavier de la requête. Il ne nous concerne pas car il est difficilement réalisable sur notre système. Il est à noter que les travaux de P. Nugues dans lesquels la parole est utilisée pour produire le textuel, permettent cette utilisation simultanée.

Le **cas n°2** fait état de deux sous-cas possibles. Les informations textuelles et déictiques sont soit en *contradiction* ou en *complément*. Dans le cas où les informations sont complémentaires, l'agent comprend que la désignation est attachée à l'objet du message et transmet le message amélioré. Le cas de la contradiction apporte deux hypothèses possibles pour l'agent.

- Hypothèse **H1** : l'agent peut supposer que l'émetteur du message désigne l'objet voulu. Il suppose que la désignation n'a pas d'ambiguïtés possible. L'agent à l'aide du moteur 3D fournit une visualisation pratique qui aide l'émetteur dans sa désignation. L'agent suppose qu'il n'y a pas d'erreur mais que sa base de connaissances est incomplète ou erronée.
- Hypothèse **H2** : l'agent suppose que l'émetteur s'est trompé et le prévient de cette ambiguïté en interrompant la transmission.

L'hypothèse **H2** a été choisie pour notre agent. Mais les possibilités apportées par l'hypothèse **H1** sont intéressantes. En effet, elle nous offre la possibilité d'enrichir les informations du monde virtuel. Le message entré par l'émetteur serait alors considéré comme une information à ajouter. Cette information fortement contextuelle serait une source d'information personnelle intéressante. Mais notre choix est resté sur l'hypothèse **H2** car les difficultés pour enregistrer cette information et l'utiliser se sont révélés complexes. Différents problèmes techniques sont apparus, notamment des problèmes de sécurité pour l'enregistrement des informations, trop éloignés de notre sujet.

Le **cas n°3** a soulevé des hypothèses similaires. Les hypothèses peuvent se réappliquer. Mais pour ce cas, nous avons choisi de ne pas offrir à l'émetteur la possibilité de désigner un objet non concerné. Car nous avons supposé les informations fournies par la modalité textuelle comme correctes. La désignation est alors considérée comme une précision ou un enrichissement apporté au texte. Une précision car le texte a pu soulever une ambiguïté ou un enrichissement car le texte était compréhensible avec l'émetteur qui a souhaité adjoindre son point de vue.

Cette analyse des différents cas d'utilisation, nous a permis de passer en revue toutes les possibilités d'interactions entre les modalités et de dégager différents modes opératoires pour agent. Mais avant de présenter les modes opératoires pour le traitement des informations, nous allons étudier le traitement des informations en sortie.

Fission multimodale

La fusion multimodale concerne la manière de traiter des informations en entrée alors que la fission multimodale concerne le traitement en sortie. Pour que les informations fournies par l'émetteur soient efficaces, il faut que l'accès à ces informations soit le plus clair possible.

Dans la théorie de la fission multimodale, l'objectif est d'améliorer la perception des informations en les distribuant sur différents canaux de perception (visuel, sonore ou haptique). Dans notre recherche, nous nous sommes concentrés sur le visuel. Certes, le visuel ne représente qu'un seul canal mais il peut être subdivisé dans notre cas en visualisation textuelle (visualisation des textes envoyés) et visualisation scénique (visualisation de la scène virtuelle). Notre recherche s'est alors portée sur la sortie la plus adaptée en fonction des informations contenues dans le message.

G. Gerbner (cf. Section 2.4.3 page 34) analyse l'importance du support pour les messages transmis. Nous savons que les communications que nous souhaitons traiter ne sont pas orientées vers un objectif défini. Nous souhaitons traiter certaines communication reconnues suivant une syntaxe. Nous étions arrivés à nous poser la question d'un support optimal en fonction d'un contenu donné (cf. Section 3.4.15 page 86). La réponse dit que dans le cadre de la communication telle que nous la concevons, le choix du canal de sortie ainsi que son emploi étaient difficiles.

Sélection par verbe

Néanmoins, nous avons cherché une approche permettant une présélection d'une des sorties d'informations. En effet, les messages que nous souhaitons obtenir sont souvent courts et leur décomposition simple. Nous avons essayé de définir une sortie possible suivant le verbe employé dans le message. Si le message contient un verbe de «mouvement» comme : déplacer, marcher. . ., nous souhaitons afficher des informations de position sur la visualisation textuelle et une visualisation scénique de l'objet concerné.

Il existe une autre solution basée sur une analyse pour chaque récepteur des différentes méthodes de visualisation choisies suivant les différents messages. Cette solution nous a semblé d'un apport peu importante vis-à-vis de notre problématique.

4.2.3 Les modes opératoires de l'agent

L'analyse précédente, nous a permis de comprendre les actions possibles suivant les différents cas d'utilisation. Maintenant, nous sommes en mesure

de définir les modes opératoires de l'ACI («Agent Conversationnel Intelligent»). Nous avons pu regrouper les actions de notre agent en cinq modes opératoires que allons expliquer.

Transmission sans traitement

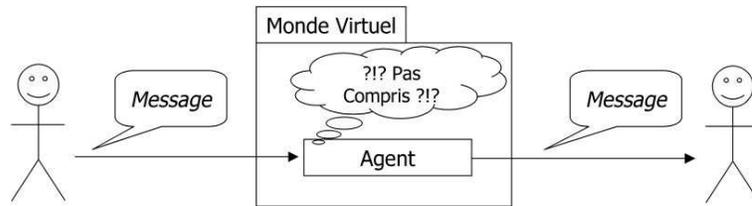


FIG. 4.7 – *Incompréhension par l'ACI*

Dans le premier cas (cf. figure 4.7), le contenu du message est incompréhensible : message trop complexe, message avec un contexte trop personnel... L'agent ne comprend pas ou n'arrive pas à déchiffrer le message. Celui-ci prévient l'émetteur de son incompréhension du message. L'émetteur peut choisir d'envoyer ou de reformuler son message.

Transmission avec traitement

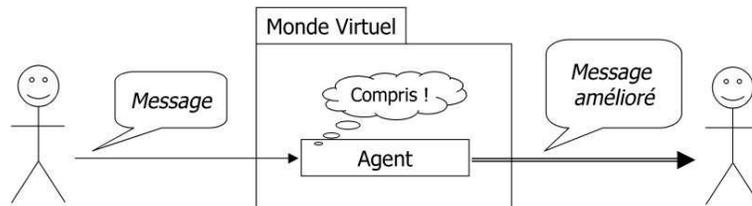


FIG. 4.8 – *Amélioration par l'ACI*

Dans le second cas (cf. figure 4.8), le message envoyé est reconnu par l'ACI. Il a réussi à analyser et à comprendre son contenu. Le message est traité, amélioré puis transmis vers le récepteur. Ce dernier reçoit le message, il peut soit visualiser le message original ou profiter de l'analyse de l'agent pour visualiser les autres informations.

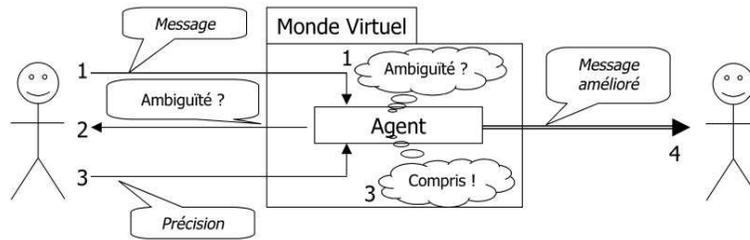


FIG. 4.9 – Levée d’ambiguïté par l’ACI

Interception et sollicitation

Dans le troisième cas (cf. figure 4.9), le message arrive à l’ACI (étape 1). Il comprend le message à transmettre, il le traite et détecte une possibilité d’ambiguïté. Il prévient l’émetteur (étape 2). Celui-ci à la possibilité de répondre aux messages de l’ACI par une désignation pour lever l’ambiguïté. Si l’émetteur décide de préciser l’objet concerné par le message (étape 3). L’ACI l’aide durant cette étape en lui fournissant des soutiens de désignation en mettant en évidence les sources de l’ambiguïté. Après désignation, l’ambiguïté est levée. L’agent procède à l’amélioration du message et sa transmission vers le récepteur (étape 4).

Interception et réponse

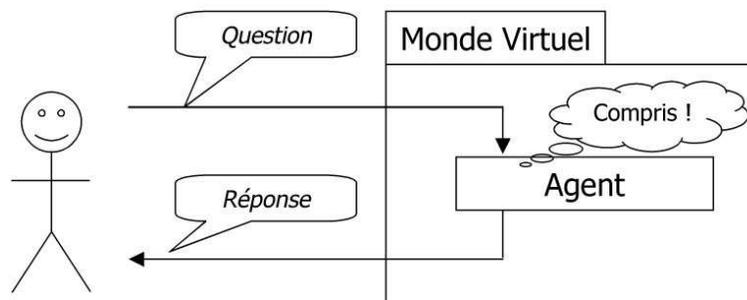


FIG. 4.10 – Réponse de l’ACI

Dans le quatrième cas (cf. figure 4.10), l’émetteur pose une question au récepteur. Si l’agent comprend le message, il peut répondre aux attentes de l’émetteur. Il propose à l’émetteur la réponse calculée.

Interception et action

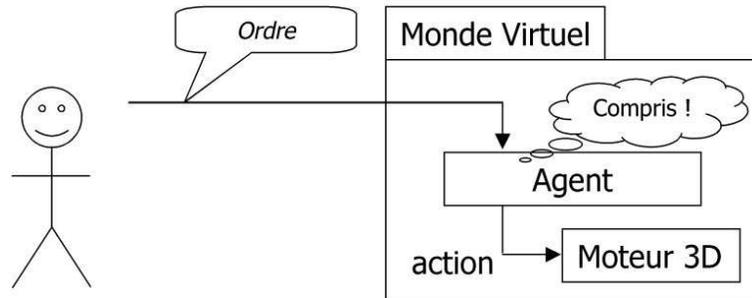


FIG. 4.11 – Action de l'ACI

Dans ce cinquième cas (cf. figure 4.11), l'émetteur envoie un ordre au récepteur. Si l'agent comprend le message, il peut effectuer l'action recherchée par l'émetteur. Il propose à l'émetteur d'effectuer l'action.

Conclusion

Tous les actes de communication transitent par l'ACI. Ce dernier va suivant le contenu du message, intervenir ou non sur les messages à transmettre. Il utilise pour cela la base de connaissances du monde virtuel. Le message déchiffré, il traite et modifie le message de l'émetteur pour adjoindre différents types d'informations.

4.2.4 Réalisation des améliorations

Nous allons présenter maintenant les réalisations effectuées pour obtenir une communication améliorée. Dans un premier temps, nous allons présenter les améliorations faites au niveau des informations du message. Nous voulions obtenir plus d'information concernant l'objet du message ; nous allons voir que les améliorations possibles peuvent se situer à différents niveaux : textuel, déictique ou visuel. Nous verrons par la suite les méthodes utilisées pour prévenir les ambiguïtés. Enfin nous expliquerons nos réalisations pour répondre aux attentes.

Étayer les informations

L'un des objectifs fondamentaux de l'agent dans la communication améliorée consiste à fournir plus d'informations. L'ACI, après l'analyse linguistique

du message à transmettre, connaît certains éléments qui le composent. Nous nous sommes concentrés sur l'objet concerné par le message et l'obtention d'information sur celui-ci.

Nous avons regroupé ces améliorations en trois catégories : textuelles, déictiques et visuelles. Nous verrons pour chacune des catégories, les améliorations que nous avons réalisées.

Les améliorations textuelles

L'objectif de ces améliorations est de fournir d'autres informations textuelles sur l'objet concerné par le message. Nous en avons conçues quatre que nous avons nommées :

1. **méthode des caractéristiques**
(cf. Chapitre 5.3.1 page 132) ;
2. **méthode de la position géographique**
(cf. Chapitre 5.3.1 page 134) ;
3. **méthode de la position visuelle**
(cf. Chapitre 5.3.1 page 137) ;
4. **méthode des fonctionnalités**
(cf. Chapitre 5.3.1 page 140) ;
5. **exploitation des méthodes textuelles**
(cf. Chapitre 5.3.1 page 141).

Reprenons l'exemple cité en introduction du mémoire concernant deux visiteurs d'un musée virtuel. (cf. problématique page 19). Lors du dialogue, les deux visiteurs souhaitent se retrouver dans le café «Chez-Pierre», mais ils n'arrivent pas à le situer. Dans la conversation, la source du problème est l'emplacement du café. Nous sommes dans un monde virtuel, l'emplacement du café dans le monde virtuel est accessible. L'agent par sa connaissance de l'objet concerné peut obtenir cette information et l'analyser. Il peut alors intercepter le message et améliorer son contenu.

Au message **Je préfère Chez-Pierre**, extrait de sa connaissance du monde des informations supplémentaires, il a la possibilité d'obtenir quatre types d'informations qui correspondent respectivement aux méthodes d'améliorations textuelles citées ci-dessus.

1. Je préfère Chez-Pierre le café avec la façade rouge.
2. Je préfère Chez-Pierre le café près de la bibliothèque.
3. Je préfère Chez-Pierre le café sur la droite.
4. Je préfère Chez-Pierre le café qui propose du café turc.

De ces quatre améliorations, il va présélectionner une amélioration avant de le transmettre (cf. figure 4.12). Cette présélection, présentée section 5.3.1 page 141, est transmise avec les autres améliorations textuelles vers le récepteur. Ce dernier pourra choisir les améliorations qu'il souhaite obtenir.

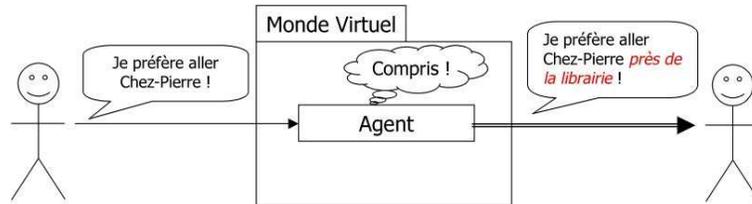


FIG. 4.12 – Exemple d'amélioration du message transmis

Les améliorations déictiques

Comme pour les améliorations textuelles, nous allons expliciter les différentes réalisations que nous avons catégorisées dans les améliorations déictiques dont voici la liste :

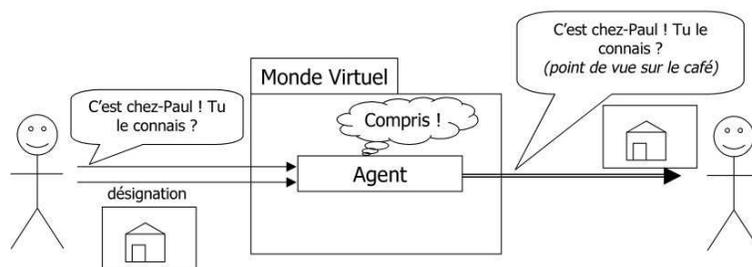
1. **désignation d'objets cachés** (cf. Chapitre 5.3.2 page 143) ;
2. **désignation avec caméra déportée** (cf. Chapitre 5.3.2 page 145) ;
3. **désignation d'ensembles d'objets** (cf. Chapitre 5.3.2 page 143) .

L'objectif de ces améliorations est de permettre à l'émetteur une plus grande expressivité en l'aidant dans la désignation. Par exemple, si nous reprenons notre scénario relatif aux visiteurs du musée à la ligne de conversation (05) (cf. problématique page 19). L'un des communicants est obligé de se déplacer pour rechercher un point de vue de l'objet. La conversation reprend seulement après cette prise d'informations. Nous proposons d'envoyer une caméra qui permet un accès plus rapide à l'objet. Ce point de vue déporté étant transmis au récepteur, cette désignation est une source d'amélioration visuelle. La communication se trouve enrichie (cf. figure 4.13).

Nous assistons là à une amélioration du support. Le texte est envoyé mais le support ne se limite plus au texte. L'agent attache au support textuel l'image utilisée lors de la désignation ainsi qu'un point vue sur la scène virtuelle.

Les améliorations visuelles

Lorsque nous considérons les travaux de G. Gerbner, nous remarquons son analyse relative à la réception des messages et à l'accès aux informa-

FIG. 4.13 – *Utilisation de la désignation améliorée*

tions. Il accorde une importance certaine à l'accès à l'information. En effet, l'accès à l'information reçue est tout aussi important que la quantité/qualité de l'émission ou de la transmission. Dans notre contexte, l'accès correspond à une visualisation de l'information. Comme le décrit aussi C. Chaillou[100], nous souhaitons dans cette partie, aider l'utilisateur dans sa visualisation et sa compréhension du monde. C'est dans cette optique qu'il apparaît important de mettre en évidence de manière explicite les objets concernés. C'est l'objet de notre troisième catégorie : les améliorations visuelles dont nous présentons la liste :

1. **mise en évidence des entités cachées**
(cf. Chapitre 5.3.3 page 148) ;
2. **transmission de point de vue**
(cf. Chapitre 5.3.3 page 147) ;
3. **visualisation par diminution d'information**
(cf. Chapitre 5.3.3 page 149) .

L'objectif de ces améliorations est de permettre une facilité d'accès aux informations par la mise en évidence des objets concernés par nos améliorations. Ainsi lorsque le récepteur reçoit un message amélioré par l'ACI, il peut demander une visualisation qu'il suppose pertinente pour lui. Dans le cas de nos visiteurs (cf. section 1 page 19), lors que A souhaite se rendre au café «Chez-Paul», E peut visualiser le café en question.

Prévenir les ambiguïtés

Une des améliorations de la communication consiste à prévenir les ambiguïtés possibles. Nous avons décidé de traiter les ambiguïtés de correspondance, c'est-à-dire entre ce qui est dit par l'émetteur et ce que l'on peut inférer du monde. La décomposition du message à transmettre permet une connaissance du contenu syntaxique. A l'aide de son contenu et surtout de

l'article utilisé pour l'objet du message, nous avons conçu deux méthodes pour prévenir des ambiguïtés. Voici la liste :

1. **prévention géographique** (cf. Chapitre 5.4.1 page 151) ;
2. **prévention visuelle** (cf. Chapitre 5.4.2 page 152).

Pour l'analyse des ambiguïtés de correspondance, nous sommes partis de l'analyse des articles (définis, indéfinis...) présents devant l'objet du message. Suivant l'article utilisé, l'agent effectue le test correspondant. Ainsi pour le cas des deux visiteurs lorsque le visiteur A envoie le message : «Je vais aller au café.». L'agent analyse l'article utilisé «au». Il lance une analyse de prévention géographique qui détecte dans notre cas une ambiguïté possible et indique :«attention, il existe plusieurs cafés.». L'émetteur prend alors conscience de l'ambiguïté pour le récepteur.

Réponse aux attentes

Dans notre approche de communication améliorée, nous avons spécifié trois actions conduisant à notre concept d'amélioration de la communication : étayer, prévenir et répondre. Ce dernier concept tente de faire intervenir l'ACI lorsque celui-ci possède les informations permettant de répondre aux attentes. Nous avons choisi de traiter deux types bien définis :

1. **réponse aux ordres** (cf. Chapitre 5.5.1 page 154) ;
2. **réponse aux questions** (cf. Chapitre 5.5.2 page 155).

Lorsque l'émetteur souhaite transmettre un ordre, l'agent analyse le verbe contenu dans le message. Cette composante renseigne sur l'action souhaitée par l'émetteur. Après analyse du monde virtuel et de l'objet concerné, l'ACI propose d'effectuer la tâche. Par exemple, si l'émetteur dit :«Emmène-moi Chez-Pierre!», l'agent comprend le souhait de l'émetteur et déplace son avatar au café «Chez-Pierre».

Concernant les questions, nous nous sommes limités à deux types de questions simples :

1. **Où est [Objet] ?** : A cette question, l'agent fournit des informations et propose d'emmener l'avatar près de l'objet concerné. Par exemple, à la question «Où est le café Chez-Pierre ?», l'agent permet une visualisation du café et plusieurs informations textuelles avant de proposer le déplacement de l'avatar de l'émetteur.
2. **Vois-tu [Objet] ?** : A cette question, l'agent analyse le contexte visuel du récepteur et offre à l'émetteur la réponse. Par la suite, l'agent

propose de montrer l'objet au récepteur. Si l'émetteur accepte, l'agent transmet un message proposant un point de vue sur l'objet concerné.

4.2.5 Conclusion

L'ensemble de ces réalisations et leur mise en œuvre via l'ACI représentent nos apports en matière de communication améliorée. L'agent conversationnel intelligent intervient dans tous les actes de communication. Il analyse le message et choisit son mode opératoire. Par la suite, l'agent utilise les différentes méthodes disponibles pour accomplir sa tâche d'aide aux actes de communication.

Les améliorations délaissées

Des améliorations imaginées, certaines n'ont pu aboutir pour diverses raisons. Nous présentons ici certaines améliorations qui auraient pu être prises en considération.

Historique des positions

Nous avons précisé dans nos écrits que les objets possèdent des caractéristiques : couleurs, forme...histoire. Nous avons exploré cette idée d'histoire des objets. Souvent l'émetteur d'un message possède des connaissances du monde virtuel à un temps T . On retrouve cette notion dans un écrit de C.Chaillou[18] sur la réalité virtuelle qui décrit la *RV* comme un support permettant d'accéder à des objets qui ont évolué dans le temps et peuvent ne plus exister ou par extension ne plus présenter les caractéristiques connues. Ainsi, en enregistrant les informations de positions dans le temps, nous pouvons désigner un objet qui a existé à un moment donné en un lieu donné. L'agent pourrait parcourir le monde dans le sens temporel inverse pour retrouver la position de l'objet voulue (ex : «prend la chaise qui était là!») Les environnements virtuels connaissent toutes les informations relatives aux objets à un temps T . Il est possible de désigner une position et de regarder «dans le temps» si «une chaise» était bien présente en ce lieu. Si nous retrouvons une chaise, rien n'empêche alors de calculer sa position actuelle.

Pour connaître la position d'un objet à un moment donné, il nous faut stocker sa position à chaque instant. Lorsque le monde comporte peu d'objets, l'enregistrement des positions est possible mais ceci se présente rarement. Un début de solution aurait été de subdiviser le monde pour enregistrer chaque élément du monde dans une subdivision et de ne sauvegarder que les

différences entre les objets virtuels. Il nous a été impossible de stocker de manière pratique et accessible toutes ces données. Nous avons donc décidé de délaisser cette possibilité.

Utilisations des sons

L'idée était de reprendre une partie des travaux que l'on peut trouver dans Diamond Park (cf. Section 3.4.7 page 64). Dans cet environnement virtuel, nous avons vu que les sons sont utilisés comme source d'informations pour renseigner les locuteurs sur le lieu où ils sont (ex : chants d'oiseaux = espace d'accueil). Ces associations de sons ont pour objectif de faire converger les locuteurs du monde vers un centre d'intérêt (ex : sons d'acclamations = centre sportif). Il apparaît alors intéressant de pouvoir utiliser les sons pour donner une ambiance et ainsi de renseigner les locuteurs présent sur un centre d'intérêt. Nous retrouvons une partie de ces idées dans VREng où un son est émis pour certaines actions notamment lorsque nous traversons un objet (ex : un son d'eau est émis lorsque nous touchons de l'eau.).

Dans l'idée de situer certains entités plus facilement, nous avons pensé émettre un son répétitif qui s'accélère à mesure que l'on s'approche de l'objet concerné par le message.

4.3 Conclusion

Le concept de communication améliorée inter-utilisateurs d'un monde virtuel constitue l'apport de cette recherche. Il regroupe un vaste domaine de cas relatifs aux différentes composantes de la communication. Nous avons choisi de nous focaliser sur le message et le médium, et de nous poser comme objectifs d'amélioration : **étayer, prévenir et répondre**. Ces apports permettent aux communications inter-utilisateurs d'être assistées par un agent qui intercepte les actes de communication afin de les améliorer avant de les transmettre au récepteur. Nous allons maintenant expliquer en détails les implémentations des méthodes au sein de notre environnement de développement «VREng».

5 Les implémentations

Sommaire

5.1	VREng : «Virtual Reality Engine»	122
5.1.1	Structure de VREng	123
5.1.2	Structure des données	124
5.1.3	Mise en place de l'agent	127
5.1.4	Conclusion	131
5.2	Implémentation des méthodes	132
5.3	Etayer les informations	132
5.3.1	Améliorations textuelles (T)	132
5.3.2	Améliorations déictiques (D)	142
5.3.3	Améliorations visuelles (V)	147
5.3.4	Conclusion sur les améliorations	151
5.4	Prévenir les ambiguïtés	151
5.4.1	Prévention des ambiguïtés géographiques (P1)	151
5.4.2	Prévention des ambiguïtés visuelles (P2)	152
5.4.3	Conclusion sur la prévention des ambiguïtés	153
5.5	Répondre aux attentes	154
5.5.1	Réponse aux ordres (R1)	154
5.5.2	Réponse aux questions (R2)	155
5.6	Conclusion	156

Dans ce chapitre, nous allons développer la mise en œuvre de la communication améliorée que nous avons présentée dans le chapitre précédent. Nous expliquerons en détails les choix effectués pour les implémenter au sein de notre support de développement. Nous débuterons ce chapitre par une présentation de «VREng», conçue et réalisée par Philippe DAX, et de quelques outils. Nous aborderons par la suite la place de l'agent au sein de cette architecture et de son fonctionnement.

Nous terminerons ce chapitre par un passage en revue des différentes méthodes mises en place qui nous ont permis d'obtenir le concept d'Agent Conversationnel Intelligent «ACI».

5.1 VREng : «Virtual Reality Engine»



FIG. 5.1 – Un musée virtuel sous VREng

VREng est une plate-forme de réalité virtuelle distribuée développée, par l'École Nationale Supérieure des Télécommunications (*ENST*) depuis 1997 dont les premiers auteurs sont : Denis Arnaud, Fabrice Bellard, Stéphane Belmon, Samuel Orzan, Lionel Ulmer[99, 5]. Actuellement, *VREng* reste disponible en Open Source sur Internet sous la license GPL. Elle permet de naviguer et d'interagir avec des objets virtuels au sein d'un monde virtuel décrit par des utilisateurs.

Les mondes virtuels sous *VREng* possèdent une grande richesse d'éléments, on peut y retrouver :

- des objets environnementaux : des sols, des murs, des planètes...
- des objets interactifs restreints : ascenseurs, portes, vagues, guides...
- des objets interactifs volatiles : des ballons, des flèches...
- des supports multimédias : clips vidéos, clips audios...
- des systèmes de particules : des étoiles filantes, de la neige...
- une représentation de notre avatar paramétrable.

Toute cette diversité tend à montrer la richesse des mondes virtuels et des possibilités du support virtuel. Cette plate-forme a été le support de nombreux travaux de recherches et de développements :

- les structures et algorithmes de données ;
- les interactions de transport ;
- les animations d'humanoïdes articulés et leurs représentations 3D avec une gestion des mouvements distribuées [Ankit Jalote, Yann Renard] ;

- création de livres interactifs [Julien Desreumaux, Sébastien Piraud] ;
- travaux sur les importations de divers modèles préconçus sous différents formats : ASE («Ascii Scene Exporter»), 3DS («3D Studiomax») [R. Piegay] , LWO («LightWave 3D Object») [P. Gros], MD2 («Quake 2 Player Model»), VNP («Vertice Normal Polygon»), OBJ [P. Dax].

Toutes ces particularités et les possibilités d'étendre les différents modules, nous ont conduit à la mise en place de notre agent et de nos méthodes d'amélioration au sein de cette plate-forme.

5.1.1 Structure de VREng

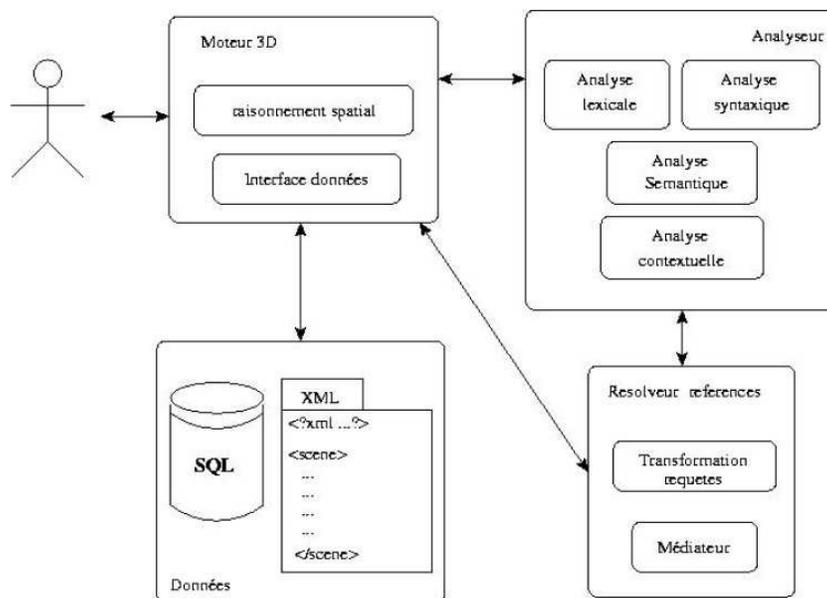


FIG. 5.2 – Schéma général de VREng

Ce premier schéma simplifié de VREng, nous présente ses composantes majeures. Il nous montre les grandes interactions entre les différents modules et les sources d'informations. En effet, au centre nous avons le moteur 3D de VREng qui gère toutes les interactions et les objets virtuels au cours de la simulation. Le second bloc en bas est la base des données. En effet, une grande partie des descriptions du monde est accessible en dehors du moteur 3D. Ces données sont stockées d'une part sous la forme de fichiers au format «vre» et d'autre part sous la forme d'une base de données. Le fichier contient des informations générales comme les noms alors que d'autre part les

positions, les orientations et autres informations plus spécifiques aux objets virtuels sont elles enregistrées dans une base de données *MySQL* mise à jour en temps-réel. C'est la composition entre les informations fichiers «vre» et le Système de Gestion des Bases de Données (*SGBD*) qui fournit à notre agent notre première base de connaissances. La seconde base est le moteur 3D lui-même.

Ces données sont importantes pour notre recherche. En effet, ces informations sur le monde virtuel permettent d'effectuer des recherches et des analyses sur les objets, leurs interactions ou encore leur contexte vis-à-vis d'un utilisateur. Il nous faut dans un premier temps rendre ces sources d'informations plus lisibles et plus faciles d'accès. La structure des données a pour but de nous fournir un outil qui nous facilitera nos développements.

5.1.2 Structure des données

Après une analyse de différents langages de description existants dans les *EV*, il nous est apparu qu'il n'existe pas de structuration universelle des données (cf. «Rapport technique sur les langages de description» [102], rapport préliminaire INRIA [43]). La structure reste adaptée aux méthodes et aux applications utilisant ces informations. En résumé, il n'est pas possible d'avoir une structure universelle pour les données. Chaque structure présente des intérêts spécifiques suivant le objectifs recherchés. Nous nous sommes dès lors tournés vers une solution que nous souhaitons pratique : XML.

Apports de XML

La structure de données est le squelette de notre agent. Il est indispensable pour obtenir une base de connaissances et d'inférences sur les objets du monde virtuel. Notre analyse de la structure des fichiers *VREng* actuels permettent de mettre en avant certains problèmes majeurs.

La structure des fichiers «vre» ne permet pas d'étendre de manière aisée les actions sur le monde. Aucune déclarations n'existent pour connaître les interactions possibles avec le monde. Les actions sont codifiées en «dur» (au cœur du moteur de simulation) et non par un langage de script comme *SCL* (cf. Section 3.4.4 page 58) ou pseudo-langage. Certes, le code source est ouvert et permet une modification du code mais le problème reste purement technique. Il nous faut extraire les actions sur les objets du code et le rendre visible au niveau de la description du *MV*. Or nous n'avons pas réussi à faire surgir seulement une partie de ces interactions et non la totalité. Dans ces conditions, il apparaît très difficile de rendre autonome cette partie de notre agent qui accède aux données (cf. caractéristiques de notre agent page 127).

Reste à voir les possibilités de rendre la description des données plus facile à utiliser.

Nous avons constaté que nous ne pouvions pas modifier une partie de la syntaxe des fichiers «vre». Nous risquions non seulement de détruire une partie des travaux mais aussi de rendre inutilisables d'autres parties. La création d'un langage est loin d'être une modification de second ordre et ne concerne qu'en partie notre recherche.

Ces difficultés nous ont orienté vers une décision de structuration par XML. Cette structuration est une étape indispensable pour la bonne marche de notre agent. Les différents avantages d'XML ont fait l'objet d'une étude dont nous présenterons les grandes lignes (cf. XML par des exemples[6, 3]). Le choix de XML s'est imposé notamment pour son lot d'utilitaires de traitements de données. Les outils de XPath pour la recherche d'un élément mais aussi XML Query Language (*XQL*) qui est un langage de requête pour XML, permettant d'interroger de manière complexe les fichiers «XMLisés» et d'en extraire des informations structurées.

Approche par XML

Nous avons aspiré à faire des fichiers de description «vre», la base de la structure d'analyse pour nos raisonnements sur le monde virtuel. Une syntaxe de document est requise pour unifier les informations décrites et vérifier l'alignement de la description sur notre modèle de syntaxe (cf. DTD <http://vreng.enst.fr/dtd/>).

Pour information, la DTD est un fichier de description syntaxique pour des contenus sous format XML. Notre première approche a été de nous diriger vers une structure standard *X3D* où les outils de visualisations et de manipulations sont nombreux. Mais ce chemin fut plus escarpé que prévu, la hiérarchisation sous le format *X3D* s'est révélée difficile. Le concept de hiérarchie n'existe pas pour le format «vre». Nous avons préféré opter vers une simplification et nous diriger vers une structuration à mi-chemin entre les formats «vre» et *X3D*.

Les fichiers «vre»

La structure résultante de la transformation en fichier XML est simple. Elle se compose en deux parties :

1. En-tête : on y retrouve des informations générales sur la version, le type de monde, les auteurs. . .
2. Scène : on y retrouve une succession de déclarations d'objets.

Les déclarations d'objet sont toutes de la même forme :

```
<type name="NOM" pos="x,y,z,ax,ay,az" dim="dx,dy,dz" ...>
  <sous-type ...>
</type>
```

Nous avons pour chaque objet, une déclaration de type qui correspond au «tag» en XML. Ce dernier permet d'affecter à l'objet, l'ensemble des fonctionnalités associées au type. Chaque type possède des sous-informations, des «propriétés» en XML comme un nom, une position, une dimension ou autre suivant le type. En effet, il existe des propriétés spécifiques (par exemple : «height» pour la hauteur des vagues pour le type «water»). De même, certains types acceptent aussi des sous-types qui ont leurs propres propriétés (par exemple le type «thing» peut accepter différentes déclaration du sous-type «solid»).

Accès par XML

Pour accéder aux éléments, nous utilisons XPath couplé avec la librairie libxml. Ainsi dans notre exemple du café, si nous supposons une description «vre» de la façon suivant :

```
<cafe name="Chez-Pierre" pos="0,1,2,3,4,5" dim="10,10,10" ...>
</cafe>
```

Lorsque nous recherchons le café «Chez-Pierre», nous utilisons comme expression de recherche xpath : `//*/[@name='Chez-Pierre']`. Nous pouvons alors accéder à l'ensemble des informations dont celui du type. Dès lors, l'agent lors de ses améliorations peut inclure dans les textes sa connaissance du type : «Je préfère Chez-Pierre, le café ...».

Conclusion

L'utilisation des fichiers de données tels qu'ils existaient à l'origine (sans une syntaxe déclarée) rendait difficile l'accès aux informations. La transformation sous un format standard nous permet de le rendre inter-opérable avec divers logiciels, permettant la lecture et la création assistées de *MV*. Néanmoins, nous sommes arrivés à un format non standard, ce qui a pour conséquence de rendre l'agent dépendant de ce format.

Notons que XQL est lié à XML, et il permet de définir un accès vers les différents éléments d'un document en XML. XPath est inspiré de XQL

et il est déjà pourvu d'une syntaxe rigoureuse. Mais il ne permet pas de mettre à jour les données, de sélectionner un groupe d'éléments et d'en extraire les informations. Notre choix s'est porté sur la `libxml` développée par GNU qui offre la capacité de parcourir l'arborescence d'un document XML, de remonter les informations demandées par l'utilisateur et fournir un nouveau document XML ne contenant que les données pertinentes. La base des connaissances se trouve ainsi aisément accessible et manipulable.

5.1.3 Mise en place de l'agent

Après avoir posé une structure pour la base des données du monde virtuel, il nous a été possible de mettre en place notre agent conversationnel. Pour ce faire, nous devons procéder par différentes étapes. D'abord, nous devons trouver une place au sein de VREng avec une interface permettant d'obtenir facilement des informations et de l'autre une interface permettant un envoi d'information et une manipulation du *MV*. C'est seulement par la suite qu'il nous a été possible de traiter les messages émis par les utilisateurs.

Architecture de l'agent

VREng est une plate-forme dont le moteur principal est écrit en C++ avec comme librairie graphique OpenGL. Divers modules coexistent dans cet *EV* suivant leur propre langage avec une interface pour interagir avec le moteur 3D. Nous avons alors décidé d'implémenter notre agent comme un module supplémentaire au sein de ces modules. Avant la mise en place de notre agent, les messages entre les utilisateurs ne subissent pas de traitements (cf. figure 5.3).

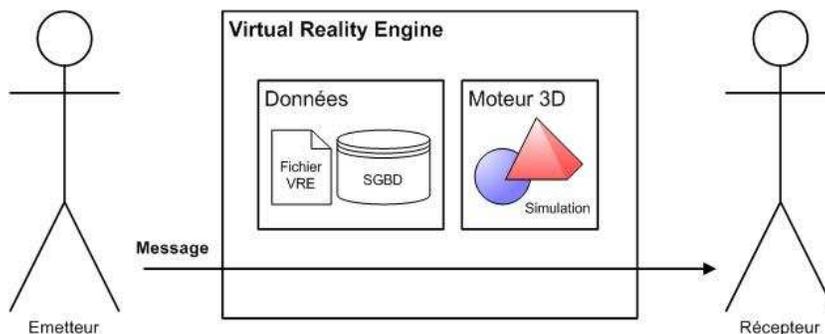


FIG. 5.3 – Schéma de VREng avant l'agent

Notre choix s'est porté sur OCaml qui permet une analyse complexe avec des moyens de générer rapidement des analyseurs lexicaux et syntaxiques (Lex et Yacc). De plus, il existe une interface permettant des liaisons entre un programme écrit en OCaml avec des programmes écrits en C++/C. Après la mise en place d'une interfaces pour recevoir des informations et une autre pour agir sur le moteur, nous nous sommes centrés sur notre agent conversationnel.

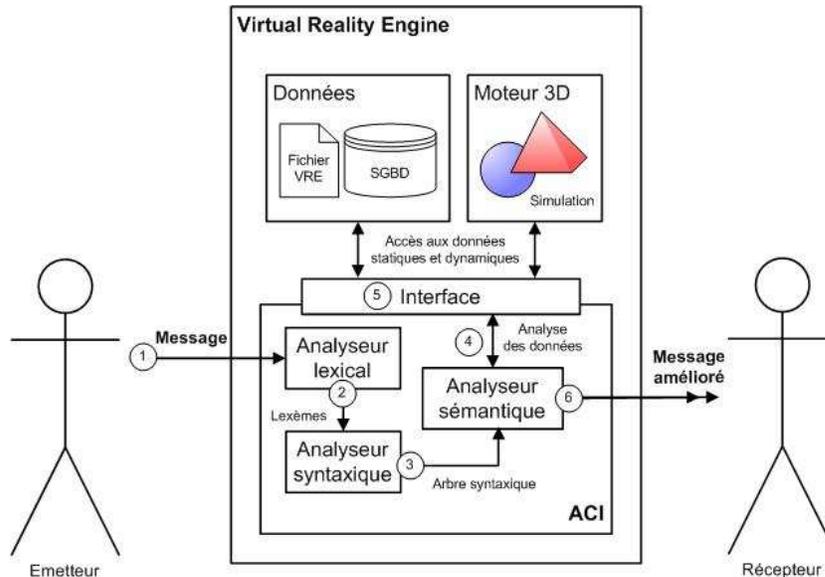


FIG. 5.4 – Schéma de VReng avec l'agent

Comment l'agent intervient sur les communications? Pour répondre, nous allons observer le schéma figure 5.4. le message textuelle émis (1) qui ne subissait jusqu'alors aucun traitement, est intercepté par le module de l'agent. Le message subit alors une série de traitements : lexical (2) et syntaxique (3). L'analyseur sémantique va alors procéder à une analyse des données (4) en utilisation l'interface de communication avec les différentes partie de la plate-forme (5). C'est au niveau du traitement (5) que l'on obtient les informations sur la désignation. L'interface va suivant les actions demandées par l'agent interroger et traiter les informations avant d'en extraire les informations exploitables. l'agent reçoit les informations (4) et les utilise pour générer un message amélioré (6).

L'agent possède tous les moyens de communication vers les différentes sources d'informations, nous allons maintenant présenter les différents analyses effectuées par l'agent pour traiter les messages textuels.

Format des messages reconnus

Un des premiers développements de notre agent s'est opéré sur la reconnaissance des messages textuels afin de comprendre les messages interceptés. Notre objectif n'est pas de reconnaître l'ensemble de tous les messages possibles. Nous nous sommes concentrés sur certains messages récurrents apparus dans l'analyse section 4.1.2 page 91.

L'analyse des messages dans les différents cas de communications médiatisées par un support virtuel nous a montré la modeste longueur des messages envoyés par l'utilisateur et leurs simplicités. Notre grammaire de reconnaissance a tenu compte de ce constat et a été simplifiée. Nous nous sommes dirigés vers la reconnaissance d'une série de types de phrases dont voici la liste :

- Message(Type, Verbe, Position, Entité, Complément) ;
- Message(Type, Verbe, Position, Entité) ;
- Message(Type, Verbe, Entité, Complément) ;
- Message(Type, Verbe, Entité) ;
- Message(Type, Verbe).

Dont voici la signification des composantes :

- **Type** : nous avons choisi de reconnaître les phrases informelles, les interrogations et les ordres à l'aide du caractère de ponctuation situé à la fin du message. Ce caractère de fin définit le type du message émis : information («.»), interrogation («?») ou ordre («!»).
- **Verbe** : Il concerne une liste de verbe. Cette reconnaissance du verbe permet une compréhension de l'action de rechercher. Il n'y a pas de module de conjugaison, seules quelques formes de conjugaison sont reconnues (le présent). Les verbes possibles sont : aller, prendre, tourner, allumer...
- **Position** : C'est une liste de positionnement relative à l'entité : en-dessous, au-dessus, à droite, à gauche...
- **Entité** : Une entité peut être de type :
 1. (Article,Objet)
 2. (Objet)

Les objets peuvent désigner soit : un **type** (ex : café), un identifiant unique (ex : «café2») ou un nom référencé dans le fichier de description (ex : Chez-Pierre) correspondant à un seul objet ou un groupe d'objets. L'article peut être : un article défini, un article indéfini...

- **Complément** : toutes les autres informations sont enregistrées dans cette partie mais ne sont pas traitées par l'agent.

Cette grammaire est réduite, de nombreux types de messages ne sont pas reconnus par notre agent mais cette grammaire regroupe une grande partie de messages inter-utilisateurs. A la suite de cette grammaire, nous nous avons pu nous atteler aux différentes analyses du message textuelle.

Analyse lexicale

Lorsqu'un message est reçu par le module agent, la première phase de traitement débute par une analyse lexicale. Cette analyse va traiter la reconnaissance lexicale du message. Si cette partie n'est pas effectuée avec succès, le traitement s'arrête avec une information sur la partie non reconnue du message. Le message n'est pas transmis et réapparaît dans le champ de saisie d'où il avait disparu quelques instants. L'utilisateur peut envoyer le message par une réitération de sa demande d'envoi auquel cas, le message n'est pas traité par l'agent car le message est identique. Sinon le message est traité à nouveau et si son analyse lexicale réussie, nous passons à la seconde analyse.

Rappel : notre choix sur le caractère de ponctuation. Lors de notre analyse des messages dans les mondes virtuels, nous avons vu que les messages envoyés étaient très courts. Les caractères de ponctuation et parfois même les caractères d'accentuations étaient omis. Ainsi l'utilisateur lorsqu'il souhaite une analyse par l'agent doit utiliser un des trois caractères reconnus : le point, le point d'exclamation et le point d'interrogation. Sans ces caractères le message ne transite pas du tout par l'analyse lexicale de l'agent. Ces caractères permettent à l'utilisateur de choisir explicitement une analyse par l'agent. Ce caractère est à la fin des messages, vu les paradigmes de «chat», il est plus facile pour l'utilisateur de retirer ou d'insérer ce caractère de ponctuation.

Analyse syntaxique

L'analyse lexicale réussie, le message passe dans le système de reconnaissance syntaxique qui va chercher à donner une structure au message. De même que pour l'analyse lexicale, le choix et les implications pour l'émetteur sont les mêmes si la reconnaissance est incorrecte. Le résultat de cette analyse va permettre l'obtention d'une arbre syntaxique respectant notre grammaire de départ. Par la suite, cette arbre est transmis à l'analyse sémantique.

Analyse sémantique

L'analyse sémantique, outre son action qui se situe avant tout sur une analyse de l'arbre syntaxique afin d'en comprendre le contenu et d'agir en conséquence, doit remplir nos objectifs : **étayer, prévenir et répondre**. La première étape consiste à étayer le message reçu afin de fournir plus d'informations sur l'objet concerné par le message textuelle. La seconde consiste à reconnaître des situations pouvant provoquer certaines ambiguïtés et ainsi lever ou du moins informer de l'existence de cette ambiguïté. La troisième consiste à répondre aux attentes de l'émetteur. A l'aide de l'arbre syntaxique l'agent va exécuter le mode opératoire choisi. Il va rechercher les informations nécessaires notamment en effectuant des requêtes *Xpath* sur les fichiers de données (cf. exemples de requêtes [6, 3]). Il va aussi accéder aux bases de données *SGBD* pour obtenir d'autres informations. Enfin, il va utilisé ses accès pour extraire les informations contenues au niveau du moteur 3D.

5.1.4 Conclusion

Notre agent (*ACI*) est capable de comprendre certains discours de l'utilisateur. Il ne comprend pas tous les messages mais lorsque le message est compris, c'est-à-dire lorsque la reconnaissance fournit des informations à l'agent. Ce dernier estime qu'il est en mesure d'intervenir suivant un des modes opératoires pour transmettre plus d'informations, prévenir les ambiguïtés ou encore répondre aux attentes de l'utilisateur.

5.2 Implémentation des méthodes

La structure et les accès aux données étant définis, nous allons maintenant présenter les implémentations des différentes réalisations citées. Il est à noter qu'une partie des explications sont purement techniques et ont été déplacées en annexe. Pour faciliter la lecture, nous avons utilisé un exemple commun, *Je préfère Chez-Pierre*. en donnant résultat obtenu pour la méthode.

5.3 Etayer les informations

Le premier objectif de notre ACI dans la communication améliorée concerne la possibilité d'offrir plus d'informations à l'utilisateur. Nous avons déjà présenté ces améliorations suivant trois catégories :

1. améliorations textuelles ;
2. améliorations déictiques ;
3. améliorations visuelles.

Nous allons reprendre cette classification pour présenter en détail les implémentations nous ayant permis d'obtenir ces améliorations.

5.3.1 Améliorations textuelles (T)

Les traitements de la modalité textuelle sont inexistantes dans les mondes virtuels et dans VREng. Nous avons pu trouver des solutions que nous allons expliquer autour d'un unique exemple.

Dans notre scénario initial, nos deux protagonistes discutent sur un lieu de rencontre. Lors de la conversation, le visiteur E souhaite transmettre une idée : *Je préfère chez-Pierre*. Si on suppose l'agent activé, il récupère le message et comprend l'objet de la conversation : *Chez-Pierre*. Il va dès lors lancer des méthodes pour obtenir des informations sur celui-ci.

A. Méthode des caractéristiques (T1)

Notre première approche a été de se pencher sur les caractéristiques des objets. Dans une conversation en général, lorsque nous évoquons un objet, nous utilisons dans de nombreux cas des informations sur l'objet et ses caractéristiques pour préciser l'objet de la conversation. De même, nous avons offert la possibilité à l'agent d'obtenir des informations sur l'objet en question. Si on applique à notre exemple, l'agent lance l'analyse des caractéristiques

et obtient en retour un pointeur vers la description XML de l'objet. L'agent analyse alors les déclarations pour en obtenir les informations.

Dans notre exemple, la description XML du café «Chez-Pierre» possède un champ indiquant l'existence d'une façade rouge. L'agent peut alors ajouter cette information au message initial :
Je préfère Chez-Pierre le café avec une façade rouge.

Explications

Nous avons expliqué que nous souhaitons une structure de données facilitant les accès. Avec cette structure en XML et une DTD déterminée, nous avons pu accéder directement à l'objet désigné par le nom «Chez-Pierre». Le pointeur retourné par la fonction de recherche de la libxml permet en outre de parcourir aisément les informations de l'objet. Ce parcours permet de choisir dans la liste des paramètres, le premier élément reconnu par l'agent, soit ici une déclaration d'un autre objet de type «facade» avec une coloration rouge.

Utilisons une partie d'un fichier de type «vre» pour illustrer les traitements effectués :

```
<mirage name= "voiture-pierre" pos="25,7,0" orientation="1.57">
<solid forme="voiture"
    taille="1.8,0.9,0.8"
    diffuse="red"
    vitesse="90" />
</mirage>

<mirage name= "voiture-jean" pos="18,13,0" orientation="3.14">
<solid shape="voiture"
    taille="1.8,0.9,0.8"
    diffuse="green" />
</mirage>
```

[issue de eric.vre]

Dans cette partie du fichier, deux voitures sont décrites. Ces deux voitures ont des positions et orientations que nous utiliserons par la suite. Les informations qui sont à notre disposition sont : des déclarations de type mirage (ces objets sont des illusions avec des propriétés liées à cet état), des déclarations de taille (connaissance de la grandeur), des déclarations de couleur (rouge, vert, etc.), des déclarations de vitesse. Notre agent va analyser les propriétés de taille et de couleur. Il va traduire la taille en texte suivant

un calcul interne par : minuscule, petit, grand ou immense. Pour les couleurs, il traduit simplement l'anglais en français. Si la déclaration des couleurs est en RGB, il calcule la distance entre les déclarations et les autres couleurs déclarées (rouge=1 0 0, jaune=1 1 0, etc.). Il choisit la plus proche pour générer le texte associé. Ainsi l'agent peut fournir des informations sur la couleur et la taille sous forme textuelle.

B. Méthode de la position géographique (T2)

Une autre méthode que l'agent peut exécuter est la méthode de la position géographique. A l'aide de cette méthode, l'agent obtient un tableau d'objets triés suivant un ordre de pertinence. Il analyse l'objet du début de ce tableau qui contient les informations de proximité par rapport à l'objet concerné par l'acte de communication. Dans notre exemple, lors de l'analyse de notre café **Chez-Pierre**, l'agent obtient comme information supplémentaire que ce dernier est proche de la **bibliothèque**.

Alors, notre agent ajoute l'information au message initial qui devient : Je préfère Chez-Pierre, le café près de la bibliothèque.

Explications

Cette «analyse géographique spatiale» permet de connaître le positionnement et les relations d'une entité virtuelle dans la scène par rapport aux autres objets. Cette connaissance permet l'ajout d'informations sur l'emplacement d'un objet. Nous allons illustrer cette analyse avec la scène de la figure 5.5.

En analysant cette scène autour d'un objet, par exemple la voiture, nous obtenons la description textuelle suivante :

- La voiture est à coté d'un arbre.
- La voiture est à coté de la maison.

Ces informations permettent de situer plus facilement et plus rapidement l'objet en question et ceci grâce à l'analyse géographique spatiale. Mais si la scène contient de nombreux objets, les informations peuvent être nombreuses et risquent de noyer le récepteur.

Si le nombre d'objets présents dans la scène est important, la description sera plus importante. Il semble évident dans ces conditions que nous effectuer une classification des informations fournies par l'analyse géographique spatiale afin de garder une clarté du message. D'où l'ajout d'un calcul dit de pertinence sur cette liste d'information.

FIG. 5.5 – *Exemple de scènes*

Pour rendre notre analyse plus pertinente, nous analysons non seulement la géographie spatiale mais aussi la taille relative. En effet, un objet éloigné présente un intérêt et un bon point de repère pour l'utilisateur s'il est de taille importante. Inversement, si un objet est très proche mais de taille négligeable, il n'est pas très intéressant d'y faire référence.

La notion de distance pure est une notion imprécise lorsqu'elle est perçue par un utilisateur. Fournir une valeur numérique n'a pas d'intérêt et ne permet pas un réel apport à la communication. Nous avons plutôt choisi de créer des classes qui regroupent les objets suivant leur proximité à l'objet concerné. Nous avons regroupé les valeurs de distance en cinq classes de distances, de la plus proche à la plus éloignée :

1. **inclus** : l'objet est inclus dans l'objet de référence.
2. **intersection** : l'objet est en intersection avec l'objet de référence.
3. **près** : l'objet est très proche, on y inclut le cas où l'objet est collé à l'objet de référence.
4. **éloigné** : l'objet en question est dans une zone éloignée.
5. **lointain** : à partir d'une certaine distance, on considère l'objet comme très éloigné de l'objet de référence.

Cette première classification permet de placer les objets dans des groupes.

Nous effectuons un traitement similaire avec la taille des objets. Tout comme la distance, la taille pure est une notion imprécise. Nous avons alors choisi de créer des classes qui regroupent les objets suivant leur taille. Le raisonnement est le même que pour les distances, nous avons classé les différentes dimensions dans cinq grandes catégories (d'une taille très petite à

celle de l'avatar humain à une taille très grande de celle de l'avatar humain)

1. **très petit** : lorsque la taille est de loin inférieure à la référence.
2. **petit** : lorsque la taille est inférieure à la référence.
3. **normal** : lorsque la taille est proche de la référence.
4. **grand** : lorsque la taille est supérieure à la référence.
5. **très grand** : lorsque la taille est de loin supérieure à la référence.

Cette seconde classification permet de placer les objets dans des groupes de taille proche.

Classification par rapport à la distance et la taille

Nous sommes en présence d'un problème de classement par rapport à un double critère. Le but est de retrouver en tête de ce classement, l'objet qu'il serait le plus pertinent à évoquer par rapport à l'objet concerné.

Pour obtenir un classement, nous avons la possibilité de trouver un classement suivant un calcul $Pertinence = \alpha(taille) + \beta(Distance)$ en effectuant un calcul de convergence pour obtenir des valeurs pour α et β .

Mais notre méthode a été différente. Nous avons pris plusieurs scènes VREng (dont Rendezvous.vre, eric.vre, maison.vre, dax.vre, etc.) et à l'aide d'un programme d'analyse, nous avons demandé de tirer 240 fois au hasard un objet de ces mondes. A chaque tirage au sort, nous avons qualifié l'objet que nous avons pensé pertinent par rapport à l'objet concerné (tiré au hasard). Le programme nous retourne un couple de valeur (taille, distance) sur l'objet désigné.

Nous reportons ce résultat dans un tableau où nous incrémentons la valeur correspondant au couple (taille, distance). Nous avons ainsi pu obtenir le tableau suivant :

	Taille 1	Taille 2	Taille 3	Taille 4	Taille 5
Distance 1	0	0	13	19	6
Distance 2	0	9	24	23	10
Distance 3	0	4	27	38	26
Distance 4	0	0	0	6	12
Distance 5	0	0	0	14	9

TAB. 5.1 – Résultat des incréments successives

A l'aide de ce tableau, nous avons pu créer un classement des couples

(taille, distance) par pertinence¹. Les valeurs varient de 1 pour peu pertinent à 25 pour très pertinent.

	Taille 1	Taille 2	Taille 3	Taille 4	Taille 5
Distance 1	3	8	18	20	13
Distance 2	5	12	21	22	19
Distance 3	4	11	24	25	23
Distance 4	2	7	10	14	17
Distance 5	1	6	9	15	16

TAB. 5.2 – Classement par pertinence

Ce classement permet de choisir dans la liste des entités virtuelles, l'objet avec la taille et la distance optimale à citer. Si plusieurs objets obtiennent le même niveau de pertinence, nous en choisissons un dans le lot qui sera celui utilisé par l'agent.

Cette méthode nous a permis un classement des informations en fonction de la taille et de la distance entre les objets. Nous nous sommes basés sur des mondes de VREng et des observations de mondes X3D disponibles en ligne pour analyser les descriptions possibles de certaines scènes. L'agent obtient pour sa part un objet avec une valeur de taille et de distance. Il va traduire en langage humain pour étayer l'acte de communication. Ainsi il ajoute à «je préfère Chez-Pierre» l'information «le café près de la bibliothèque».

C. Méthode de la position visuelle (T3)

Une autre méthode que l'agent peut utiliser est celle du calcul de la position visuelle. Nous avons vu à plusieurs reprises que l'utilisation du contexte virtuel du récepteur serait un avantage certain pour la compréhension de certains actes de communication. Pour notre exemple : «Je préfère Chez-Pierre», si l'agent fait appel à cette méthode, il obtient trois éléments : *des coordonnées, un objet et un pourcentage*. Les *coordonnées* représentent la position de l'objet à l'écran, le *pourcentage* est le pourcentage d'occultation de l'objet et *l'objet*, l'objet qui occulte visuellement l'objet concerné.

Ainsi lorsque l'agent analyse ces informations disponibles, il peut améliorer le texte par : Je préfère Chez-Pierre, le café sur la droite.

¹Ce classement tient compte pour chaque couple (taille, distance) des valeurs alentours, après plusieurs essais, nous sommes arrivés au tableau de classement par pertinence

Explications

Cette méthode «d'analyse de la position visuelle» permet d'obtenir une information sur la perception relative de l'utilisateur concernant la scène. Suivant le point de vue que l'on a de la scène les informations de positions sont différentes. Illustrons ce problème par la scène de la figure 5.6.

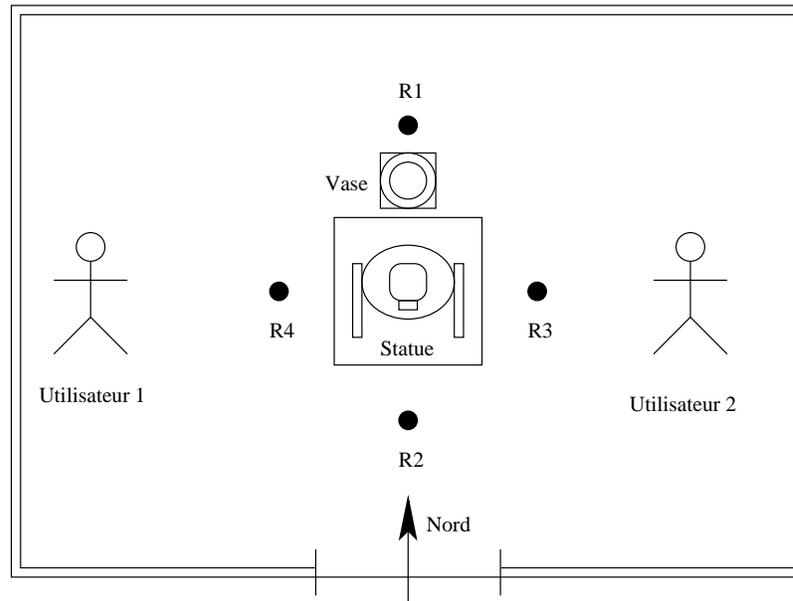


FIG. 5.6 – *Illustration de la perception relative au positionnement*

La flèche du bas représente l'entrée de la pièce. Nous avons schématisé une scène usuelle dans un musée où une statue se situe au centre de la pièce. Les avatars des utilisateurs déambulent et se retrouvent de part et d'autre de la statue. A cet instant, le contexte visuel est différent pour les utilisateurs. les points **R_i** représentent les positions d'observation du vase et de la statue.

Le «vase» possède une position relative différente selon le point de vue des utilisateurs : le vase se situe à gauche de la statue pour l'utilisateur 1 (**R₄**) et à droite de la statue pour l'utilisateur 2 (**R₃**).

Un article de Pfeiffer[76] montre qu'il existe d'autres points de vue : le point de vue déporté (associé à un objet par exemple **R₁**) ou le point de vue absolu (associé au monde **R₂**). Si on analyse cette même scène suivant ces nouveaux points de vue, on obtient que le vase est derrière ou devant la statue. Les travaux de Pfeiffer montre que le choix de la meilleure référence dépend de nombreux contextes.

Dans le cadre de notre communication interpersonnelle, les messages sont

destinés au récepteur. Nous avons alors choisi de traiter la position visuelle selon le contexte visuel du récepteur. En effet, il est plus facile pour le récepteur du message de comprendre le contenu d'un message lorsqu'il fournit des informations suivant son propre point de vue. Lorsque l'émetteur transmet le message au récepteur, l'agent peut demander une analyse visuelle du monde virtuel dans le contexte visuel de récepteur pour obtenir les informations de position.

Le contexte visuel de l'utilisateur peut être reproduit par le moteur en arrière-plan de l'application à l'aide des Buffers annexes (cf. off-screen rendering, Annexe 8.4). Avec l'accès au contexte visuel du récepteur, il est possible de calculer la position d'un objet sur l'écran du récepteur. Par l'application d'une fonction de projection OpenGL, on obtient alors des coordonnées (x, y) qui correspondent à la position de l'objet sur la fenêtre de visualisation. L'agent utilise ces coordonnées et les informations de la fenêtre de visualisation pour calculer la position de celui-ci. L'agent va générer un texte suivant le cas détecté. Si l'objet est visible à droite, il génère le texte «à droite» ; si l'objet n'est pas à l'écran mais sur la droite, il génère le texte «sur la droite». Par exemple, le message «Je préfère Chez-Pierre» devient «Je préfère Chez-Pierre, le café sur la droite»

Cas des objets occultés

Il n'est pas rare de voir les objets partiellement ou totalement occultés. Nous avons alors cherché à informer le récepteur du message que l'objet présente cet état de fait. Dans ce but, nous avons cherché à connaître le pourcentage de surface occultée et l'objet en premier plan de cette occultation.

Il n'existe pas de méthode directe permettant ce calcul. Certes, nous avons eu connaissance de l'une des fonctions en cours de développement pour les nouvelles cartes graphiques et utilisable par la librairie OpenGL : la fonction «glOccluded». Cette fonction semble prometteuse lorsqu'elle sera accessible.

Nous avons cherché d'autres méthodes permettant d'obtenir ce résultat. En annexe sous le titre «Analyse du mode de picking» (cf. Annexe 8.2 page 177). On y détaille deux méthodes permettant d'obtenir les objets situés sous des coordonnées (x, y) suivant le mode de «picking». La discussion concernant le mode de picking est technique, sa compréhension est laissée à titre informel.

Notre fonction de calcul a pour but de calculer le nombre de pixels visibles de l'objet (en premier plan) NbV sur le nombre de pixels de l'objet (premier plan ou occulté) $NbNV$. Le calcul du pourcentage se résume en un calcul

entre le rapport entre NbV et $NbNV$.

En suivant le raisonnement précédent, nous avons souhaité employer la méthode de picking sur tous les pixels à l'écran afin de connaître le pourcentage d'occultation. Le résultat est optimal, nous obtenons toutes les informations utiles : le nombre de pixels visibles, ceux occultés, les objets qui l'occultent. Malheureusement ce calcul a été impossible à mettre en place au sein de VREng pour des raisons de ressources graphiques et CPU.

La solution adoptée a été de réduire les calculs. Pour calculer le pourcentage, nous calculons 5 points de projection de l'objet. Nous effectuons un picking sur ces 5 points. Suivant les résultats, le pourcentage varie de 0 à 100% par étape de 20%. Et l'objet ayant le plus de fois occulté lors du calcul est considéré comme l'objet occultant l'objet sujet de la conversation.

Le résultat de notre calcul correspond dans la plupart des cas à nos attentes. Cette méthode nous semble efficace notamment parce que le monde de VREng est surtout composé de parallélépipèdes et non de formes trop complexes, impliquant des positions simples. La méthode renvoie un pourcentage et un objet `Obj`. L'agent emploie ces informations pour ajouter au message textuel l'information : «derrière [`Obj`]» si le pourcentage dépasse les 50%.

D. Méthode des fonctionnalités (T4)

Regardons cette nouvelle méthodes : celles des fonctionnalités. En effet, dans notre recherche de désignation textuelle d'une entité virtuelle, nous avons cherché ce qui discriminait un objet d'un autre. La notion de fonctionnalité nous a semblé comme un axe possible d'amélioration. Les fonctionnalités d'un objet nous renseignent sur sa nature et précise l'objet. On constate que l'énonciation d'une fonctionnalité nous permet de discriminer un objet dans un ensemble. Dans notre exemple, l'agent peut utiliser cette méthode pour obtenir des informations sur les fonctionnalités de l'objet. En retour de cette méthode, l'agent reçoit une liste de fonctionnalités associées à son type et une liste des fonctionnalités associées à l'objet précis. L'agent va calculer la différence entre ces deux ensembles. Il obtient alors une liste de fonctionnalités propre à cet objet.

L'ACI utilise une fonctionnalité qu'il possède dans sa liste interne pour convertir cette fonctionnalité en texte. Il va ainsi traduire notre message d'exemple en : je préfère Chez-Pierre, le café qui propose du café turc.
--

Explications

Lors du lancement initial de VREng, chaque type d'objet déclare toutes les fonctions voulues. Ces différentes fonctions sont nommées et répertoriées par le moteur. Ainsi, il maintient une table sur toutes les fonctions possibles. Lorsque nous connaissons une entité virtuelle, nous pouvons donc connaître toutes ses fonctionnalités en analysant le type de l'objet.

La première partie de notre amélioration au niveau de la fonctionnalité a été de fournir à l'agent toutes les fonctionnalités de l'objet concerné. Ainsi l'agent peut exprimer des différences fondamentales. Par exemple, si l'objet concerné est une porte, l'agent ajoute «qui peut s'ouvrir» ou dans le cas d'une lampe «qui peut s'allumer»).

Ceci fait, nous nous sommes concentré sur la seconde partie de notre amélioration qui est encore en développement : nous avons voulu exprimer des différences plus fines entre plusieurs mêmes objets. Nous souhaitons agir au niveau du code XML en ajoutant ou retirant des fonctionnalités. Ceci nous servira à donner aux objets des comportements précis pour qu'ils se différencient d'autres objets de même type. Ainsi nous pouvons différencier les cafés par leurs spécificités. On peut ainsi différencier les cafés qui font du thé turc des autres cafés. La manière d'obtenir la liste des fonctionnalités de l'objet s'effectue à l'aide de la librairie de parcours XML qui retourne les déclarations fonctionnelles de l'objet et en produit une liste de fonctionnalités nommées.

Ces fonctionnalités sont décrites par un verbe qui est traduit par l'agent pour générer un texte explicatif de type : «qui [fonctionnalité]». Par exemple, pour le message «Je préfère Chez-Pierre», l'analyse des fonctionnalités du café «Chez-Pierre» montre que la fonction «faire café turc» est propre à ce café. L'agent peut alors fournir un texte amélioré «Je préfère Chez-Pierre, le café qui propose du café turc».

E. Exploitation des méthodes textuelles (T5)

Nous avons vu qu'il existait différentes méthodes pour améliorer les informations textuelles. Nous cherchons maintenant à faciliter la visualisation de ces messages textuels : le récepteur reçoit ces améliorations et visualise l'amélioration suggérée par l'agent. Il a la possibilité de visualiser les autres améliorations possibles ou de visualiser le message original de l'émetteur.

Explications

Lorsque le récepteur reçoit le message, il reçoit non pas un message mais une liste de cinq messages et un chiffre. Dans cette liste, on retrouve en première position le message original suivi des autres messages textuels améliorés par chacune de nos méthodes.

L'interface reçoit la liste et le numéro du message à faire paraître en premier. Il affiche sur l'interface de dialogue le message avec l'indice du numéro. A la vue du message, l'utilisateur peut visualiser les autres améliorations à l'aide du système de menu personnalisable (M. Leboïte). Ce menu déroulant liste le nom des différentes méthodes et la sélection dans la liste de l'une d'elles provoque l'affichage de l'amélioration demandée.

Nous avons souhaité que l'agent propose directement une amélioration sans une demande explicite. Dans ce but, nous avons choisi d'analyser le verbe employé par l'utilisateur pour connaître l'amélioration la plus appropriée. Le verbe permet la description de l'action, il est l'information que nous avons pensé la plus pertinente pour ce choix. L'agent possède une liste de verbes reconnus avec des valeurs. Ces valeurs définissent le choix de l'amélioration première à faire paraître au récepteur.

En conclusion

A l'aide de ces méthodes, l'agent va pouvoir étayer les informations sous la forme textuelle qu'il peut ajouter au message initial. L'agent exécute une par une les méthodes, il obtient alors des améliorations possibles que le récepteur pourra visualiser s'il le souhaite.

5.3.2 Améliorations déictiques (D)

Nous avons employé la modalité déictique comme une modalité secondaire rattachée à la modalité textuelle. Notre amélioration de la modalité déictique s'est alors centrée sur l'élargissement des conditions d'entrée. En d'autres termes, nous avons recherché à améliorer l'expressivité et donc améliorer la communication.

Nous allons présenter ici différentes méthodes qui ont permis d'améliorer la désignation en offrant à l'utilisateur des outils «facilitant» la désignation.

A. Désignation d'ensemble (D1)

Par cette méthode, nous avons voulu offrir la possibilité de désignation d'un ensemble. Nous connaissons deux types de désignation d'ensemble : la sélection un à un et l'effet «lasso». Rapidement, cette méthode consiste à appuyer sur une touche et sélectionner un par un les objets, on forme ainsi un groupe de plus en plus grand. La seconde désignation consiste à décrire un encadrement pour sélectionner tous les objets se trouvant à l'intérieur.

Pour notre part, nous avons cherché une nouvelle méthode basée sur la proximité. Dans les améliorations textuelles, nous avons vu la méthode de la position géographique **T2**. Elle permet de fournir une liste triée des objets suivant la proximité par rapport à un objet. Nous avons cherché à l'appliquer pour la désignation d'un ensemble.

Explications

Le principe de cette méthode repose sur le fait que la première sélection s'effectue sur l'objet dit de référence. Cette sélection effectuée, nous lançons dans le menu la demande de sélection d'un ensemble. Nous utilisons une partie de l'amélioration **T2**. Le moteur sélectionne tous les objets ayant un niveau de proximité 1. L'utilisateur peut augmenter ou diminuer le niveau, ce qui a pour effet d'agrandir ou de diminuer les niveaux sélectionnés. La sélection du niveau n sélectionne les niveaux inférieurs. L'utilisateur peut ainsi inclure non seulement des objets visibles mais aussi les objets non-visibles ou difficiles à sélectionner. Par cet emploi on sélectionne des blocs continus d'objets.

Pour améliorer encore cette désignation, nous avons ajouté la possibilité de sélectionner un ensemble homogène. Par homogène, on entend des objets de même type. Par exemple, sur un bureau si on souhaite sélectionner seulement les livres et non les autres objets, on lance la sélection d'un ensemble, on choisit la désignation d'un ensemble homogène. On observe alors une sélection qui varie seulement au niveau des livres. (cf. figure 5.7).

B. Désignation d'objets cachés (D2)

Dans l'optique d'élargir les conditions d'entrée de la désignation, nous avons cherché à améliorer la communication en augmentant l'expressivité de l'utilisateur. Après la désignation d'un ensemble, nous avons cherché à désigner des objets virtuels partiellement ou complètement occultés.

Lors des actes de désignation, il n'est pas rare de voir des objets partiellement occultés. Nous proposons de les désigner au travers des objets



FIG. 5.7 – Désignation d'un ensemble de livres

occultants. La désignation traverse les objets occultants et désigne l'objet voulu.

Explications

Pour ce faire, nous utilisons un principe OpenGL que nous avons vu précédemment pour la méthode de la position visuelle (**T3**) : le «picking». Mais la méthode de la position visuelle est différente de ce que nous souhaitons faire.

Le mode « picking » est parfaitement adapté à nos besoins et cette désignation nécessite peu de moyens pour être mise en place. En effet, si on simplifie la vision des principes d'affichage, on peut considérer le monde visualisé à l'écran comme une superposition en couches des différents objets du monde virtuel sur notre plan de visualisation (cf. méthode «picking» Annexe 8.2 page 177).

Lors d'une sélection, la sélection s'opère à l'aide de cette méthode qui permet de connaître à l'endroit de désignation : le nombre d'objets superposés (nombre de hits) et la liste des objets superposés.

Lors de la sélection, nous choisissons le premier objet visible comme la plupart des méthodes de désignation. Mais nous offrons la possibilité de sélectionner les objets placés «sous» la couche principale. Cette méthode est facilitée par OpenGL qui trie la liste des objets superposés suivant la «profondeur» des objets et donc des couches. L'utilisateur n'a plus alors qu'à faire défiler la sélection des objets un par un. L'utilisateur appelle cette

fonction à l'aide de la souris ou d'un menu qui sélectionne successivement les objets présents sous le dernier point désigné.

C. Désignation avec caméra déportée

Toujours pour permettre à l'utilisateur une meilleure expressivité et ainsi améliorer la communication, nous avons mis au point une méthode de désignation avec une caméra déportée. Nous sommes partis du constat que nous devons très souvent nous déplacer et nous positionner pour désigner. Nous avons cherché à faciliter ces déplacements par une caméra qui effectue ce déplacement. L'utilisateur peut rester à la même place pour observer la scène. Cette technique d'utilisation de caméra a pour objectif d'assister la désignation par une navigation. Suivant les travaux de L. Nigay [72], cette navigation implique un objectif à atteindre comme : trouver un objet ou un ensemble d'objets de la scène. Afin de parfaire notre technique, nous avons essayé de respecter les règles énoncées par L. Nigay visant une bonne navigation. Ces règles consistent à faciliter le changement de visualisation en assurant la continuité visuelle lors du passage d'une visualisation à une autre.

Ainsi dans le scénario d'introduction, le visiteur A qui souhaite sélectionner le café n'a plus besoin de se déplacer. Il lance une caméra qui va se déplacer et proposer un point de vue sur l'objet souhaité.

Explications

Cette méthode consiste à avoir dans la fenêtre de visualisation principale, une seconde fenêtre de visualisation sur le monde. Cette seconde visualisation consiste à offrir un mode sélection par étapes. Nous avons pensé utiliser une visualisation directement dans le monde comme dans les travaux de Lenay[56] où la vision du point de vue des autres utilisateurs est visible sur le visage des utilisateurs. Pour des raisons simples, cette méthode ne permet pas une sélection aisée pour l'utilisateur.

Notre méthode utilise des options de commandes OpenGL «`glScissor`» et «`glViewport`» qui nous permettent respectivement de dessiner et d'afficher sur une partie de la fenêtre de visualisation. La manipulation des paramètres de cette commande permet de gérer la zone d'affichage. Nous pouvons ainsi simuler sur une partie seulement de la zone d'affichage, une autre vision du même monde. Un second point de vue est alors disponible (cf. figure 5.8).

Le protocole d'utilisation de cette caméra se déroule par étapes. En premier, l'utilisateur sélectionne un objet ou un ensemble d'objets vers lequel il souhaite envoyer la caméra. Une caméra est alors envoyée vers une position



FIG. 5.8 – *Visualisation du point de vue de la caméra*

optimale calculée automatiquement (cf. Annexe «Positionnement automatique» page 8.3). A ce moment là, l'utilisateur de la caméra peut effectuer deux actions : sélectionner ou déplacer la caméra.

Ouvrons une parenthèse pour rappeler que VREng possède deux types de désignation : le premier, modélisé par un trait continu, permet la sélection d'un objet ; le second, modélisé par un trait discontinu, permet la désignation d'une direction.

Nous utilisons ces différents traits pour différencier la demande de sélection et la demande de déplacement. Au niveau de la visualisation secondaire (la caméra), si on sélectionne à l'aide d'un trait continu un objet alors l'utilisation de la caméra se termine. Si l'utilisateur désigne à l'aide du trait discontinu la caméra va se déplacer. Elle va adopter un autre point de vue en suivant la demande de l'utilisateur. Elle ira se placer sur un autre objet ou offrir un autre point de vue si l'utilisateur désigne le même objet.

En conclusion

Notre apport au travers de ces désignations a été d'améliorer l'expressivité de l'utilisateur. Nous avons élargi les conditions d'entrée de la modalité. La modalité déictique a été améliorée ce qui permet suivant notre définition d'améliorer la communication. La désignation était jusque là limitée aux objets visibles seuls. Nous pouvons maintenant désigner des blocs d'objets homogènes ou hétérogènes et cela sans qu'ils soient visibles.

5.3.3 Améliorations visuelles (V)

Les améliorations en entrée (amélioration en entrée) nous ont permis d'obtenir des informations supplémentaires. L'amélioration des informations en sortie (amélioration visuelles) a pour objectif de proposer une visualisation améliorée. L'objectif est de permettre à l'utilisateur d'accéder à l'information améliorée pour les visualiser.

A. Transmission de point de vue (V1)

Dans les *EV* existants la transmission d'un message est souvent textuelle. Or, il est parfois difficile de décrire textuellement des situations ou des actions. Il est préférable de les montrer. Avec cette idée, nous avons cherché à améliorer le support du message en ajoutant une image au message textuel. Quand l'agent a compris le message, il va distinguer deux cas : utilisation ou non de la désignation.

Sans utilisation de la désignation, l'agent connaît l'objet de la conversation : il va rechercher dans l'ensemble des objets virtuels de la scène l'objet correspondant. Il va alors prendre une «photo» de la scène et l'envoyer en attaché au message textuel.

Avec utilisation de la désignation, l'agent possède non seulement l'objet de la conversation mais aussi un point de vue que l'émetteur a utilisé pour désigner l'objet. Nous sommes partis du principe que l'émetteur pouvait envoyer le point de vue de la désignation comme image de l'objet. L'agent a connaissance des désignations. Il va prendre ce point de vue et l'envoyer au récepteur.

Si nous reprenons le scénario principal A (ligne 01) : «**Je t'attends au café Chez-Paul**», l'agent prend connaissance de la position de l'objet, crée une image du café «**Chez-Paul**». Le récepteur reçoit le message avec l'objet «**Chez-Paul**» sous la forme d'un lien hypertexte qui montre une image du café. Si l'utilisateur fait un double-click sur ce lien, un point de vue secondaire est alors créé comme pour la caméra déportée (section 5.3.2 page 145). Le récepteur peut ainsi visualiser l'objet directement en temps-réel.

Explications

L'implémentation de cette fonctionnalité a été soumise à des problèmes techniques. Dans cette section, nous allons présenter l'implémentation de cette fonctionnalité au sein de la plateforme VREng pour laquelle un problème de ressources réseau s'est présenté. Nous n'avions pas les moyens de transmettre directement les images au récepteur. Les explications techniques

sont disponible en annexe (cf. Annexe 8.4 page 182).

Nous avons donc essayé de trouver un autre moyen de transmettre la un point de vue. Pour palier à ce problème, rappelons un constat des travaux de Gerbner (cf. Chapitre 2.4.3 page 34) sur une différence fondamentale de la réalité virtuelle et la communication réelle : l'environnement virtuel crée un contexte commun. Dès lors, il est intéressant de transmettre non pas des images mais seulement des références. Ou plus exactement des coordonnées dans la scène qui correspondent à l'image voulue. L'interface se charge alors de traiter l'information et de générer l'image au niveau du récepteur.

La méthode en place est simple, l'agent va générer des informations de coordonnées et l'associer à l'objet. Par exemple, si nous reprenons notre exemple : «Je t'attends au café Chez-Paul.» va être modifié par l'agent en : « Je t'attends au café @ x, y, z, ax, ay, az @Chez-Pierre@. »

x, y, z sont les coordonnées spatiales absolues et ax, ay, az les orientations. Le message est reçu en amont avant la visualisation de l'utilisateur. L'interface retire les coordonnées du message reçu et fait générer une image. La génération de l'image est faite suivant la méthode dite de « off-screen rendering » (pour les explications cf. Annexe 8.4 page 182).

Enfin pour obtenir un second point de vue, nous affectons au lien hypertexte la possibilité de lancer une caméra déportée aux coordonnées indiquées (cf. figure 5.9). L'émetteur transmet non seulement un texte mais attaché à ce texte des coordonnées qui sont transformées en image et point de vue.



FIG. 5.9 – Visualisation d'un point de vue

B. Mise en évidence (D2)

La transmission du point du vue nous semble être une méthode utile et efficace pour désigner un objet mais son emploi est parfois limité selon les contextes. Les objets peuvent être trop éloignés, occultés ou encore au milieu d'autres objets similaires.

Nous nous sommes penchés sur la mise en évidence des objets pour faciliter la visualisation de l'objet. Dans les *MV* que nous avons étudiés, les mises en évidence sont pour le moins simplistes. Elles se limitent à un ajout d'un encadrement en surimpression autour de l'objet. L'utilisateur qui observe la scène voit alors l'objet entouré par des traits de sélection. Après la mise en place de méthodes similaires : flèches de direction ou encadrement à densité de traits variables, nous avons trouvé que les problèmes pour situer l'objet était résolu mais pas la visualisation de l'objet en lui-même. Nous avons alors mis en place deux méthodes qui apportent leurs modes de visualisation : la méthode de surimpression (utilisation d'un effet technique d'OpenGL) et la méthode de diminution des informations (utilisation d'un concept de visualisation).

Surimpression

Ce paradigme consiste à mettre en évidence un objet par une visualisation au travers de toutes les autres objets. Pour ce faire nous utilisons une méthode de surimpression (notamment par un emploi de la fonction de transparence OpenGL `GL_BLEND`). L'effet obtenu est permet une visualisation de la position de l'objet mais aussi de la nature de l'objet. Cet effet est obtenu par la visualisation en monocouleur de sa texture externe.

Cette méthode montre son intérêt lorsque la différence entre la surface de projection en mode normal et en mode surimpression est forte. On entend par là que son application est efficace lorsque l'objet de taille grande et fortement occultée. Elle a moins d'intérêt lorsque l'objet est de petite taille ou lorsque l'objet concerné est très éloigné. En effet, dans ces conditions, les informations de textures deviennent inutiles.

Diminution d'informations

Dans de nombreux domaines, la recherche d'objets tend à une augmentation des informations visuelles pour améliorer la compréhension. Nous sommes partis d'un raisonnement inverse pour arriver à cette méthode de visualisation. Nous avons cherché à diminuer la quantité d'informations visuelles afin de mettre en évidence l'objet concerné.

Comme on le voit sur la Figure 5.10, nous affichons en mode «fil de fer» tous les objets non-concernés et nous laissons un affichage «complet» de l'objet concerné. Dans la pratique, nous avons trouvé son emploi utile lorsque l'objet concerné était entouré d'autres objets avec des caractéristiques similaires. Il est à noter que dans certains cas, cette visualisation ne nous permet pas de situer l'objet dans la scène notamment lorsque qu'il y a de nombreux

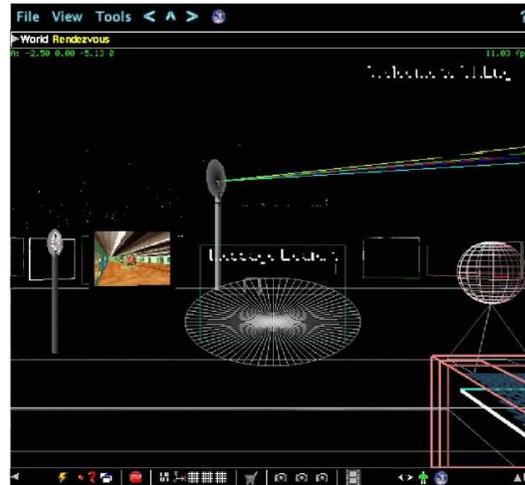


FIG. 5.10 – *Visualisation du point de vue de la caméra*

objets ou lorsqu'il est très éloigné.

Explications

La mise en évidence d'un objet est disponible à l'aide du menu personnalisable de VREng où les fonctions peuvent être lancées. Le lancement de ces méthodes est soumis à la condition qu'un objet doit déjà être sélectionné. Lorsqu'un message est transmis avec un objet concerné, l'utilisateur doit pouvoir lancer directement une mise en évidence de celui-ci. Pour résoudre ce problème, nous avons utilisé une méthode similaire à la transmission de point de vue. Nous ajoutons à la demande de point de vue, le nom unique de l'objet dans le monde virtuel. Lors d'un acte de communication, l'interface réceptionne le message, il trouve l'objet concerné à l'aide de son identifiant et le sélectionne avant d'effectuer d'autres traitements.

Lorsque le message initial est : «Je préfère Chez-Pierre.»,
le message résultant de cette méthode est : « Je préfère
@x,y,z,ax,ay,az@cafe – num2@Chez-Pierre@. »

En conclusion

Au travers de ces améliorations en sortie, nous avons cherché des moyens d'aider l'utilisateur à accéder à l'information. G. Gerbner accorde une grande importance à la compréhension des messages. Nous avons voulu offrir à l'utilisateur diverses solutions et de moyens de contrôle sur l'information. Dans

certaines conditions, nous proposons l'aide de solutions de visualisation que nous pensons être efficace. L'utilisateur reste maître des améliorations auxquelles il accède. Par exemple, la transmission du point de vue reste constamment accessible et seul une volonté du récepteur permet sa visualisation.

5.3.4 Conclusion sur les améliorations

L'ensemble des ces méthodes permettent à l'agent de remplir l'objectif d'étayer les communications. L'emploi des améliorations textuelles fournissent de nombreuses informations. Les améliorations déictiques permettent une plus grande expressivité à l'émetteur. Le récepteur d'un message peut aisément accéder à ces informations. Nous avons réussi à améliorer la communication en enrichissant le **contenu du message** et le **support du message**.

5.4 Prévenir les ambiguïtés

Un autre objectif de notre amélioration de la communication, a été la mise en place de différentes méthodes permettant la prévention d'ambiguïtés. Sans pouvoir traiter tous les cas d'ambiguïtés, nous avons cherché à traiter deux cas précis : les ambiguïtés géographiques et les ambiguïtés visuelles.

5.4.1 Prévention des ambiguïtés géographiques (P1)

L'objectif de notre méthode a été la mise en place d'une méthode permettant de détecter des incohérences linguistiques vis-à-vis du contexte géographique de la scène. Dans les conversations usuelles, nous avons parfois une connaissance précise et une méconnaissance de ce qui existe. Ainsi une évidence peut faire apparaître des ambiguïtés. Par exemple, si nous reprenons notre scénario avec nos deux interlocuteurs du musée virtuel. L'un des premiers messages dit : « **Je vais t'attendre au café.** ». S'il existe dans les environs différents cafés, il apparaît évident que cette phrase peut induire une ambiguïté. Il faudrait pouvoir prévenir et demander à l'utilisateur une précision. L'agent lorsqu'il reçoit le message, lance une méthode de prévention géographique. Dans ce cas, il lance la méthode avec les coordonnées de l'utilisateur. En retour, la méthode fournit une liste des objets de type «café».

L'agent analyse cette liste, si elle contient plus d'un élément, il prévient éventuellement de l'existence de plusieurs cafés dans la zone de l'utilisateur. «**Il existe plusieurs cafés : Chez-Pierre, Chez-Paul, Chez-Jacques. lequel ?**»

Explications de l'algorithme

Le lancement de la méthode de prévention n'est pas effectué pour tous les actes de communication. Lors de l'émission du message, celui-ci est décomposé en plusieurs éléments dont la composante «Objet» qui désigne l'objet concerné. Cette décomposition permet aussi d'obtenir l'article de l'objet. Si l'article est un article défini singulier : «le», «la», «l'», «au» et «du», l'agent déclenche le processus de prévention.

Le processus reprend une partie de la méthode d'analyse géographique (**T2**) (cf. section 5.3.1 page 134). L'agent utilise les coordonnées de l'avatar comme objet de référence et lance une analyse géographique. Une liste des objets de même type est produite après élimination des objets «très éloignés» (critère de distance). L'agent peut ainsi générer un message de prévention suivant la taille de la liste :

- taille < 1 : «Il n'existe a pas de [type]».
- taille = 1 : aucune ambiguïté détectée, le traitement continue. L'agent effectue une modification de la décomposition du message. L'objet reconnu seulement par son type devient alors le café «Chez-Paul». L'objet concerné étant parfaitement connu, le message peut subir les autres traitements de l'agent.
- taille > 1 : «Il existe plusieurs [type]». L'agent propose une utilisation du support déictique pour préciser l'objet concerné.

Lorsque la taille est supérieure à 1, nous avons choisi d'ajouter des informations sur les objets à l'origine de l'ambiguïté. Lorsque nous obtenons la liste des objets, nous pouvons accéder à leurs informations. Dans notre exemple, il y a trois cafés dans les environs avec comme nom : «Chez-Pierre», «Chez-Paul» et «Chez-Jacques». Nous reprenons ces noms et nous les ajoutons au message de prévention des ambiguïtés.

5.4.2 Prévention des ambiguïtés visuelles (P2)

Alors que la première méthode a pour but la prévention des ambiguïtés géographiques, celle que nous allons présenter a pour but d'éviter certaines ambiguïtés visuelles. Par exemple, si nous reprenons le scénario et modifions la requête de l'émetteur par : «Je vais t'attendre à ce café». Nous pouvons vérifier au niveau de l'émetteur s'il n'y a pas une ambiguïté visuelle : présence d'un autre café. L'agent lance alors la méthode de prévention visuelle qui lui retourne une liste d'objets. Suivant la taille de cette liste, l'agent va générer ou non un message.

Dans notre exemple, si on suppose que l'émetteur était en face de deux cafés, l'agent prévient en envoyant le message : «**Il existe plusieurs cafés devant vous : Chez-Pierre, Chez-Paul, Chez-Jacques**»

Explications

Alors que la prévention géographique traite suivant les articles définis, cette prévention s'effectue lorsque des adjectifs démonstratifs singuliers sont détectés : «ce», «cet» ou «cette».

Dans ce cas, l'agent lance la méthode de prévention géographique et obtient une liste des objets de même type. Pour chacun des objets de cette liste, l'agent lance le calcul de la position visuelle de l'objet à l'écran. Ce calcul va permettre de savoir si ces objets ont une chance ou non d'être visibles sur l'écran de l'utilisateur. A partir de ce calcul, on élimine de la liste ceux qui n'ont aucune chance d'être visibles à l'écran. On calcule la visibilité de chacun plus précisément avec l'appel de la méthode de la position visuelle (**T3**) (cf. section 5.3.1 page 137). La méthode s'arrête si plus d'une entité de même type est visible.

En retour, l'agent reçoit une liste d'éléments. Cette liste regroupe les d'entités visibles appartenant au type concerné (ex : café). Suivant la taille de la liste, l'agent va ainsi générer différents textes :

- taille < 1 : «**Je ne vois aucun [type].**»
- taille = 1 : L'existence d'une seule entité permet à l'agent de connaître l'objet précis concerné par l'acte de communication. Le café n'est plus «un» café mais le café «Chez-Pierre». L'objet concerné connu, l'agent peut alors effectuer les autres traitements.
- taille > 1 : «**Je vois plusieurs [type] :**», l'agent propose une désignation pour lever l'ambiguïté.

Tout comme la prévention géographique, nous avons les moyens de connaître à l'aide de cette liste, les autres cafés visibles. Nous pouvons dès lors informer l'utilisateur de l'existence des autres cafés lorsqu'il en détecte plusieurs.

Par exemple, l'agent peut générer un texte et prévenir l'émetteur : «**Il y a plusieurs cafés devant vous : Chez-Pierre, Chez-Paul, Chez-Jacques**»

5.4.3 Conclusion sur la prévention des ambiguïtés

L'agent peut utiliser deux informations du monde virtuel dans lequel l'émetteur émet un message. Soit il utilise la position géographique des ob-

jets, soit la visualisation de l'émetteur. Par ces deux informations, nous pouvons prévenir de deux ambiguïtés possibles. De plus, les méthodes employées restent proches des autres traitements qui nous ont permis une implémentation rapide et simplifiée de cette approche.

5.5 Répondre aux attentes

Le dernier objectif fixé pour l'amélioration de la communication concerne la possibilité de répondre aux attentes de l'émetteur. Nous cherchons à extraire des informations ou à agir sur le monde. Dans cette partie, nous allons expliquer les deux types d'attentes de l'émetteur d'un message que nous avons souhaité traiter : répondre aux ordres et répondre aux questions.

5.5.1 Réponse aux ordres (R1)

L'émetteur peut souhaiter faire une action sur le monde. Dans ces conditions, nous avons voulu que l'agent puisse intervenir et proposer d'effectuer l'ordre.

Ainsi lorsque l'émetteur donne un ordre d'action tel que : «**Ouvre la porte principale !**», l'agent peut comprendre le message et agir en conséquence. Ainsi l'agent peut proposer le déplacement de l'émetteur vers l'objet voulu.

Explications

La décomposition du message émis par les analyseurs nous permet d'obtenir le «verbe» de la phrase. Nous allons utiliser les informations du Verbe pour connaître l'action recherchée. L'agent prend connaissance du verbe et de l'existence d'un objet dans le message. L'agent peut se retrouver dans deux situations possibles :

1. l'objet du message existe ;
2. l'objet du message est inexistant.

Dans le premier cas, nous recherchons l'objet dans la liste des objets qui composent le monde virtuel. L'objet trouvé, nous effectuons un traitement similaire à l'analyse des fonctionnalités (**T4**) (cf. section 5.3.1 page 140), nous analysons les fonctionnalités disponibles à ce type d'objet. Si une des fonctionnalités de l'objet correspond à la demande de l'utilisateur, l'agent propose de l'effectuer. Ainsi si l'émetteur dit : «**Ouvre la porte principale !**», l'agent accède à l'objet, il constate l'existence de la fonction «ouvrir» et propose d'effectuer l'action.

Dans le second cas, l'agent va rechercher dans la liste des objets visibles s'il existe un et un seul objet possédant cette fonction. S'il trouve un seul et unique objet qui possède cette fonctionnalité, l'agent propose d'effectuer l'action.

Par exemple, si l'émetteur voit une seule porte et s'exprime par : «Ouvre !». L'agent comprend et effectue l'action sur la porte.

5.5.2 Réponse aux questions (R2)

Il n'est pas rare de s'interroger sur des informations sur le monde. Nous avons souhaité que ces interrogations sur le monde aient directement une réponse de l'agent. Ainsi lorsque l'émetteur souhaite connaître des informations sur un objet du monde, nous pouvons lui fournir différentes informations sur celui-ci.

Par exemple, si l'émetteur ne connaît pas l'emplacement d'un objet, il peut s'interroger sur la position de celui-ci : «Où est le café Chez-Pierre?». L'agent peut ainsi intervenir et répondre sur l'emplacement du café et offrir un point de vue sur celui-ci.

Explications

Nous avons choisi de traiter deux types de questions :

1. les questions de position ;
2. les questions de visualisation.

Pour les questions de position, l'agent traite les messages textuels de la forme : «Où est [Objet]?». L'émetteur a désigné l'objet et souhaite obtenir des informations de positions sur celui-ci. L'agent cherche l'objet dans les éléments du monde et utilise les différentes améliorations textuelles et visuelles pour informer l'émetteur. L'agent va générer un message avec les différentes améliorations textuelles, il va ajouter en attaché une image et un point de vue de l'objet.

Ainsi, l'agent peut répondre à la question «Où est le café Chez-Pierre?» en une phrase de la forme : «Le café Chez-Pierre près de la fontaine [image de l'objet].» et si l'objet est à une distance éloignée, il propose de déplacer l'avatar vers ce café.

Pour les questions de visualisations, nous avons voulu que l'agent puisse traiter les questions sur le contexte visuel du récepteur. L'émetteur souhaite connaître un information visuelle de l'utilisateur et demande : «Vois-tu [Objet]?». L'agent comprend l'objet concerné par le message et lance une analyse

du contexte visuel du récepteur. L'agent utilise la méthode d'analyse visuelle (**T3**) sur l'objet concerné et reçoit les informations permettant de répondre soit :

- «Oui, il le voit [image du contexte du récepteur] .», s'il est à l'écran et si le pourcentage visible est supérieur à 50%.
- «Non, il est caché par [Objet] .», s'il est à l'écran mais avec un pourcentage inférieur à 50%.
- «Non, il est sur sa [gauche/droite] .», s'il est tout juste hors de l'écran de visualisation.
- «Non, il est derrière lui .», s'il est à l'opposé de la visualisation.

L'agent utilise le contexte du récepteur pour calculer la visibilité de l'objet. Les résultats de ce calcul fournissent à l'agent les informations pour générer des informations

L'analyse visuelle est une réutilisation de la méthode de position visuelle. L'agent qui analyse la position visuelle à l'aide de la position de l'émetteur, va utiliser celui du récepteur. L'information obtenue permet à l'agent de répondre précisément à la question de l'émetteur.

Conclusion sur la réponse aux attentes

Par ces méthodes, l'émetteur peut ainsi interroger ou agir sur les éléments du monde virtuel. L'agent traite différents types d'ordres liés aux fonctionnalités disponibles sur les objets virtuels. Il intervient lorsque qu'il a les compétences requises pour satisfaire l'ordre émis. De même, il a la capacité de répondre à des questions sur le positionnement ou la visualisation des objets virtuels.

5.6 Conclusion

Notre agent ACI et ses méthodes ont été implémenté au sein de la plateforme VREng. Ils forment l'ensemble de nos apports pour le concept de communication améliorée. L'agent introduit au sein de VREng, analyse et intervient dans les actes de communication.

Une des difficultés notable liée aux modes virtuels est l'accès aux informations. L'environnement détient l'ensemble des informations de simulation de la scène mais ces mêmes informations sont disséminées. Nous avons dû parfois nous plonger au cœur du moteur pour extraire des informations pertinentes et les rendre utilisables.

Notre modèle d'agent n'est pas aussi généraliste que nous l'aurions sou-

haité. Les modalités traitées sont usuelles et présentent un format d'information standard mais nos implémentations tiennent compte des propriétés de la plate-forme virtuelle. Tous les environnements virtuels ne sont pas propices à l'élaboration et à l'emploi de notre agent et de nos méthodes. L'architecture du monde est une source d'information importante et de nombreux paramètres sont directement liés à sa structure et à son architecture. Cette constatation rend les améliorations fortement liées aux possibilités de l'environnement virtuel. Cependant, nous avons tenté de rendre indépendant (modulaire) les composants de notre agent. Nous avons séparé et regroupé certains traitements. D'une part, nous avons les possibilités d'accès aux données statiques (fichiers), dynamiques (SGBD) et temporaires (moteur 3D) et d'autre part, nous avons les analyses possibles destinées au traitement des messages textuels.

Si certaines méthodes sont fortement liées au moteur de simulation 3D, l'agent lui présente une certaine indépendance. La refonte des différents interfaces est inévitable, elle accomplit la liaison entre l'agent et le moteur de simulation. Nous pensons que les apports de cet agent avec les théories de la communication améliorée, nous ont ouvert des possibilités d'élargissements de certaines notions comme la communication assistée.

6 Evaluation

Sommaire

6.1	Principes de Newcomb	159
6.2	Scénarios sur la communication améliorée	160
6.2.1	Scénario sans ACI : S1	160
6.2.2	Scénario avec ACI : S2	161
6.2.3	Conclusion	163
6.3	Proposition pour l'évaluation de l'ACI	164
6.4	Prototype d'ACI	165

Après l'analyse détaillée des implémentations visant une amélioration des actes de communication, nous arrivons à la dernière partie de ce manuscrit : l'évaluation. Il nous faut évaluer ces différentes méthodes et la réalisation de notre Agent Conversationnel Intelligent. Après une explication par l'exemple de notre méthode d'analyse, nous présentons un protocole permettant d'évaluer notre agent.

6.1 Principes de Newcomb

Pour quantifier notre amélioration de la communication, nous allons décrire un protocole expérimental permettant de mesurer cet apport. Une communication améliorée est une communication où les ambiguïtés sont levées et l'acte de communication fournit à l'utilisateur des informations claires et pertinentes. Suite aux travaux de Newcomb (cf. Partie 2.4.4, page 37), nous retenons que la communication peut être vue comme une séquence d'actes avec un objectif de convergence vers une situation (ex : se retrouver dans un café). Si nous reprenons cette modélisation de la communication, notre concept de communication améliorée peut être vu comme une diminution du nombre d'échanges pour atteindre cette situation. Les locuteurs peuvent ainsi atteindre cette situation plus rapidement. C'est à partir de ce principe que nous avons conçu l'évaluation de nos méthodes d'amélioration de la communication.

6.2 Scénarios sur la communication améliorée

Dans cette partie, nous allons présenter les différentes améliorations à travers un scénario. Si nos deux visiteurs de l'exposition virtuelle (cf. Chapitre 1, page 19) avaient eu à leur disposition un agent conversationnel intelligent, quelle aurait été le fil du dialogue ?

6.2.1 Scénario sans ACI : S1

Dans l'introduction de ce mémoire nous avons présenté un dialogue entre deux visiteurs pour se rejoindre dans un café après la visite d'une exposition virtuelle.

Contexte

- Cadre : visite d'une exposition virtuelle d'œuvres d'art ;
- Visiteurs : A et E ;
- Support de communication : boîte de dialogue (*chat*) et outil de désignation au sein d'un monde virtuel.

A a terminé la visite. E est encore à l'exposition.

A souhaite retrouver E dans un cadre convivial.

A se dirige vers un café pour attendre E.

Une conversation s'engage entre A et E :

01. A -> E : J'ai terminé la visite; je t'attends au café.
02. E -> A : D'accord, celui qui est à côté de la fontaine ?
03. A -> E : Non, celui qui est près de la bibliothèque.
04. E -> A : Je ne vois pas.
05. A -> E : Attends, je vais voir son nom.
 {déplacement de A vers le café}
 Il s'appelle: Chez-Paul ! Tu le connais ?
06. E -> A : Non, ce nom ne me dit rien.
07. A -> E : Tu préfères aller dans un autre café ?
08. E -> A : Oui, je préfère Chez-Pierre.
09. A -> E : Où est le café Chez-Pierre ?
10. E -> A : Il est près de la fontaine,
 tu ne peux pas le rater.
11. A -> E : D'accord, j'y vais.

Légende :

- «X -> Y : M» signifie X envoie le message M à Y ;
- «{Action}» signifie Action commandée par l'utilisateur et effectuée par l'avatar de l'utilisateur.

Analyse

Détaillons la situation, A et E sont dans des zones différentes. A souhaite indiquer un lieu de rendez-vous à E. En retour, E cherche à connaître le lieu de rendez-vous pour y arriver. A et E s'échangent des informations pour arriver à l'état de «se retrouver dans un lieu convivial».

Dans ce scénario **S1**, les deux parties se sont échangés **11 messages** pour réussir à trouver le lieu de rencontre commun. Si on analyse le dialogue nous y retrouvons de nombreuses imperfections du langage. Lors du premier échange, une ambiguïté apparaît (01) (Il existe plusieurs cafés). E connaît les cafés des environs mais n'arrive pas à situer celui-ci (02, 03). Ceci oblige E à interroger A afin de lever l'ambiguïté sur le café (04). Mais A est dans l'incapacité d'y répondre directement et doit se déplacer afin de trouver une information pour préciser l'entité concerné (05). Le café en question semble déplaire à E (06). Les interlocuteurs choisissent de se fixer un rendez-vous dans un autre café (07, 08) mais A ne peut le situer correctement. E souhaite transmettre ses connaissances sur la position. Il transmet alors les informations qu'il pense utiles pour A le récepteur (10). A décide alors de s'y rendre (11).

Ce scénario met en exergue différentes difficultés actuelles de la communication médiatisée par un support virtuel. Les ambiguïtés persistent, les informations sont parfois difficiles à transmettre. Le support textuel est parfois difficile d'emploi sous certaines conditions.

6.2.2 Scénario avec ACI : S2

Si nous reconsidérons le précédent scénario avec l'hypothèse que deux visiteurs de monde virtuel ont à leur disposition un agent conversationnel intelligent (ACI) tel que nous l'avons décrit. Nous pouvons supposer que la conversation se retrouve modifiée ainsi :

Contexte

- Cadre : visite d'une exposition virtuelle d'œuvres d'art.
- Visiteurs : A et E
- Agent conversationnel intelligent (*ACI*)

- Support de communication : boîte de dialogue (*chat*) et outil de désignation analysée par ACI
- Communication entre A et E : tutoiement.
- Communication entre A ou E et ACI : vouvoiement.

A a terminé la visite. E est encore à l'exposition.

A souhaite retrouver E dans un cadre convivial.

A se dirige vers un café pour attendre E.

Une conversation s'engage entre A - ACI - E :

01. A -> E (ACI) : J'ai terminé la visite;
je t'attends au café.
02. ACI -> A : Il y a plusieurs cafés à proximité :
Chez-Pierre, Chez-Paul, Chez-Jacques.
Précisez ?
03. A -> ACI : <pointe vers le café>.
04. ACI -> A : C'est Chez-Paul; je vous montre.
05. {point de vue de A:
face au café Chez-Paul}
06. A -> ACI : C'est bien celui-ci.
- 07* A (ACI) -> E : J'ai fini la visite; je t'attends au café
Chez-Paul près de la bibliothèque.
08. E -> A (ACI) : {E demande un point de vue du café}
- 09* E -> A : Non, il n'est pas très engageant,
je préfère Chez-Pierre.
10. A -> E (ACI) : Où est le café Chez-Pierre ?
11. ACI -> A : Le café Chez-Pierre est sur la droite
près de la fontaine.
Je vous emmène ?
12. A -> ACI : {A active le déplacement vers le café}
- 13* A -> E : D'accord, j'y vais.

Légende :

- «X -> Y : M» signifie X envoie le message M à Y ;
- «X -> Y(ACI) : M» signifie X veut envoyer le message M à Y mais il est intercepté et compris par l'ACI ;
- « <désignation> » signifie action de désignation ;
- «{Action}» signifie Action effectuée par l'avatar de l'utilisateur ;
- «*» signifie indicateur d'un message inter-utilisateurs.

Analyse

Nous nous situons dans le même contexte avec les mêmes supports de communication que lors du scénario S1. La différence se situe au niveau des messages transmis qui sont analysés par notre ACI. L'agent raisonne sur la base des informations du message et des informations relatives au monde virtuel.

Dans ce scénario S2, le résultat final est identique, les utilisateurs ont réussi à se donner un rendez-vous dans un café. Un premier constat pour ce scénario concerne le nombre de messages entre A et E qui n'est que de **3 échanges** (07, 09, 13) contre **11 échanges** pour le scénario S1.

Dès le premier message, l'ACI détecte l'existence d'une ambiguïté (02) à cause des différents cafés présents aux alentours. L'agent mentionne à A la source de cette ambiguïté et attend une information complémentaire (03) permettant une levée de l'ambiguïté. A ne connaissant pas le nom du café, il précise la position de ce dernier et utilise la caméra déportée pour faciliter la désignation du café en question (03). Après confirmation, l'ACI comprend et améliore le message initial avec un point de vue sur le café (07). Après réception du message amélioré, E observe le point de vue attaché et fait une proposition à A. Ce message n'est pas compris par l'agent qui le transmet sans traitement (09). A cherche à interroger E sur l'emplacement de ce café (10). La question de A est comprise par l'ACI. Ce dernier propose de répondre à la question de A et de l'y emmener (11). A accepte le déplacement et confirme le lieu de rendez-vous (13).

Un second constat concerne le nombre total d'échanges qui est plus important dans le scénario S2. Si le nombre d'échanges entre les locuteurs a diminué, le nombre de messages a augmenté. Nous pouvons expliquer cela par le fait que l'agent évite certains actes entre locuteurs. Il n'a pas pour objectif de diminuer l'ensemble des actes. Il a pour objectif de diminuer les actes entre les locuteurs.

6.2.3 Conclusion

A travers ces scénarios (S1 et S2) avec et sans ACI, nous voyons l'intérêt de notre agent conversationnel intelligent à assister les communicants. Nous observons un effet quantifiable déjà décrit dans les travaux de Newcomb : le nombre de messages échangés. La communication passe par différentes étapes pour arriver à un point de convergence. Ce nombre de messages échangés dépend directement des informations contenues dans les messages transmis et des attentes des utilisateurs. L'objectif de nos améliorations est d'augmenter ces informations contenues dans les messages transmis.

Ces améliorations devraient avoir pour conséquences directes de diminuer le nombre de messages échangés.

6.3 Proposition pour l'évaluation de l'ACI

Nous avons mis en place des protocoles pour comprendre et évaluer les différentes méthodes d'améliorations.

Les propriétés

Comme nous l'avons vu précédemment, nous pouvons utiliser le nombre d'échanges inter-utilisateurs pour évaluer nos méthodes. Pour évaluer l'apport possible de ces méthodes dans les actes de communication pour pouvons recourir à deux tests d'utilisation :

1. un test témoin où le support et les informations ne sont pas améliorés
2. un test où l'ACI est utilisé. Les actes de communications sont surveillés par l'agent en vue d'une amélioration possible.

Chaque test et chaque scénario devront utiliser deux sujets. Il est préférable dans un premier temps d'utiliser seulement un communicant humain qui teste l'amélioration. Le second communicant effectuera le protocole en connaissant les améliorations possibles de l'agent. Ces deux utilisateurs devront effectuer une activité collaborative où les échanges seront enregistrés ainsi que le durée de l'ensemble du scénario depuis son lancement.

Description de l'activité envisagée

Le scénario envisagé consiste en une activité d'aménagement d'une pièce avec différents éléments comme : un sofa, une table, des chaises, etc. Un plan de rangement est disponible à l'extérieur de la pièce. L'objectif est de laisser les déménageurs virtuels communiquer pour aménager la pièce. Certaines actions s'effectueront sans échange. Le but est de laisser les communicants agir. Nous souhaitons retrouver dans les actes, des messages reconnus par l'agent, tels que : «Où est le sofa ?», «Prend cette chaise!». Nous pourrions alors vérifier par rapport au scénario sans agent, les apports de l'agent lors de ses interventions.

Conclusion de ce protocole

Le premier résultat proviendra du nombre d'échanges entre les utilisateurs. Le résultat pourra se présenter sous la forme d'un tableau à deux entrées avec d'un côté le nombres de messages entre les utilisateurs sans agent et de l'autre, le nombre d'échanges en excluant les interventions de l'agent. Il nous faut aussi tenir compte du temps. En effet, les utilisateurs pourraient avoir tendance à ne pas communiquer et faisant peu usage des avantages de l'agent. Il serait intéressant d'observer les répercussions de ces avantages en fonction du temps. Par la suite, d'autres informations pourront être étudiées comme le nombre d'échanges entre les utilisateurs et l'agent.

Résultats attendus

Les résultats attendus sont une diminution du nombre de messages échangés entre les utilisateurs lors de l'utilisation de l'ACI. Le protocole requiert un temps important. Nous pensons que ce protocole permet une évaluation de notre système et de notre approche de la communication améliorée par un ACI.

6.4 Prototype d'ACI

Dans le cadre de l'évaluation des principes de notre ACI, un démonstrateur a été programmé pour nous permettre de présenter quelques exemples d'améliorations que nous avons décrits dans les méthodes d'amélioration. Ce démonstrateur se focalise sur les possibilités de répondre aux attentes de l'utilisateur.

Analyse des fonctionnalités

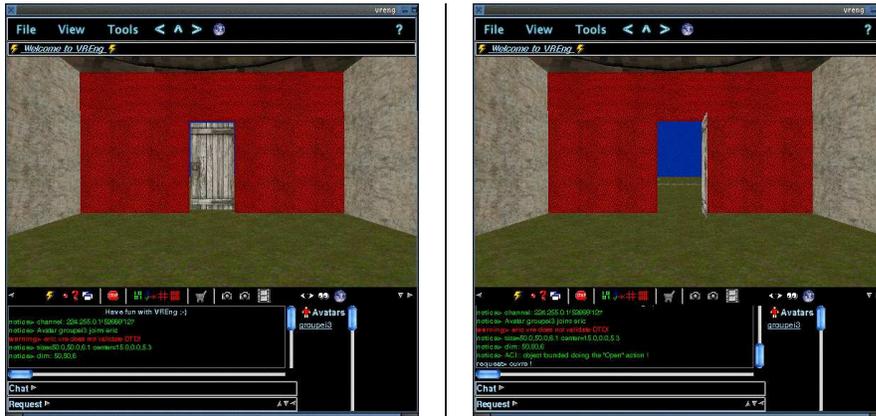


FIG. 6.1 – *Utilisateur : «Ouvre !»*

Dans cet exemple (cf. Figure 6.1), l'utilisateur est à l'intérieur d'une maison. Il s'approche de la porte de sortie et émet un souhait : «Ouvre!». L'ACI intercepte le message et l'analyse. L'agent détecte dans le contexte visuel de l'utilisateur un objet de type «porte» qui possède la fonctionnalité de s'ouvrir. Cet objet est le seul qui peut remplir cette condition, l'agent va alors décider d'ouvrir la porte à l'utilisateur.

Analyse du contexte visuel



FIG. 6.2 – *Utilisateur : «Ferme cette porte !»*

Dans cette scène (cf. Figure 6.2), l'utilisateur est sorti d'une maison et souhaite fermer la porte de celle-ci. Il émet un souhait précis, celui de fer-

mer une porte qu'il désigne par l'article démonstratif «Ferme **cette** porte!». L'analyse visuelle montre qu'un seul objet de type «porte» est visible devant l'utilisateur. L'agent prend alors la décision de répondre aux attentes de l'utilisateur en fermant cette porte.

Analyse du contexte visuel pour détecter des ambiguïtés

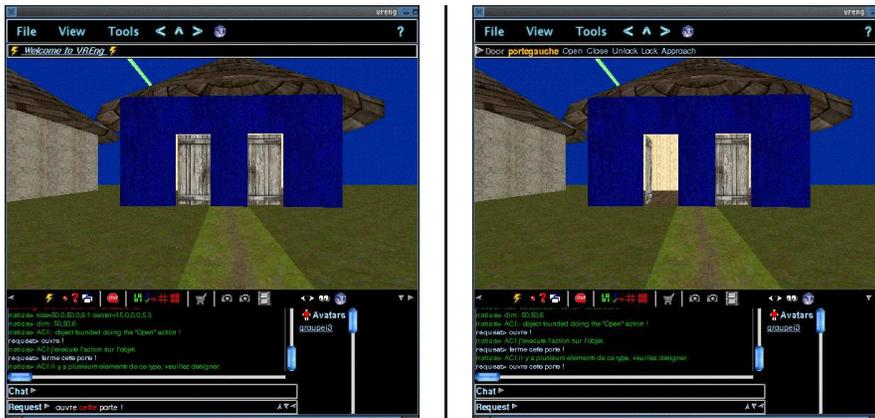


FIG. 6.3 – Utilisateur : «Ouvre cette porte!»

Dans ce dernier exemple (cf. Figure 6.3), l'utilisateur émet le souhait d'ouvrir une porte précise qu'il désigne au travers de l'article «cette». Or l'agent analyse le contexte visuel de l'utilisateur et détecte la présence de deux portes qui peuvent être une source d'ambiguïtés. L'ACI prévient l'utilisateur et lui demande de préciser sa requête. L'utilisateur désigne alors la porte concernée en la sélectionnant à l'écran à l'aide de la souris. L'agent possède dès lors toutes les informations pour agir et ouvre la porte qui a été désignée.

7 Conclusions et perspectives

Conclusions

Ce travail s'inscrit dans le contexte de recherches relatives au domaine de la réalité virtuelle. Il concerne la communication entre humains connectés à un même monde virtuel. L'objectif est d'étudier les apports de ce nouveau contexte. Notre problématique peut être résumée par deux questions (page 5) auxquelles nous avons essayé de répondre.

Question 1 : *Quelles sont les caractéristiques spécifiques de la **communication** dans un **monde virtuel** ?*

Le travail relatif à la communication nous a permis d'appréhender les problématiques récurrentes à ce domaine. Nous avons pu en extraire les principaux modèles qui nous ont permis de comprendre certains formalismes comme les décompositions de Lasswell ou l'analyse du contexte de Gerbner.

L'état de l'art nous a conduits à confronter ces formalismes au domaine de la réalité virtuelle et nous a permis de comprendre les possibilités d'un tel domaine en termes d'outil, de médium pour la communication. La réalité virtuelle nous est apparue comme un support de conversation innovant.

Question 2 : *Quelles **améliorations** de la communication sont réalisables grâce au **virtuel** ?*

Les apports résultant du contexte réalité virtuelle concernent deux aspects :

- les moyens de communication ; nous mettons en œuvre les modalités : textuelle, déictique, ainsi que leur conjonction ;
- les contenus des échanges ; nous avons introduit un agent conversationnel intelligent doté d'une connaissance du monde virtuel et capable

d'interagir avec l'utilisateur

Ces apports ont conduit au concept de **communication améliorée** relative à une amélioration d'une ou de plusieurs composantes de la communication. Nous nous sommes concentrés sur deux de ces composantes : **le contenu et le support**.

Concernant le contenu, la réalisation d'un tel concept nous a conduits à l'élaboration d'un **Agent Conversationnel Intelligent**. Il a pour but de réaliser les objectifs de la communication suivant différents modes opératoires comme l'interception du message dans le but de l'améliorer ou encore l'anticipation d'une ambiguïté dans le message, suivie d'une interrogation de l'émetteur pour sa levée.

Concernant le support, la communication médiatisée par un environnement virtuel offre de nombreux avantages. La connaissance des paramètres qui régissent la simulation est l'un d'entre eux. En outre, cette connaissance permet une prise en compte des contextes propres à chaque utilisateur du monde virtuel, action difficile, voire impossible dans le réel (exemple : l'analyse du contexte visuel d'un autre utilisateur). Il existe toutefois des limitations liées à ce support. Son champ d'action ne couvre pas toutes les possibilités de la communication telles que le principe du «Chat». Cela se traduit par une faiblesse de l'expressivité des utilisateurs face à des mondes virtuels souvent formatés pour des tâches précises. Nombreux sont les mondes comme Habbo (page 71) qui permettent une grande richesse de la communication textuelle sans offrir la possibilité de transmettre une désignation par pointage.

Ainsi, nous avons pu apporter quelques éléments de réponse aux deux questions initiales. Par cette étude, nous avons pu mesurer l'ampleur des possibilités d'améliorations accessibles grâce au virtuel. La difficulté de leur réalisation réside dans la capacité à extraire les informations du virtuel et inférer de ces informations des améliorations pertinentes. L'environnement virtuel offre de larges possibilités d'innovation dans le domaine de la communication. L'environnement virtuel a permis d'agir sur deux des composantes de la communication : le contenu et le support. La manipulation des informations (exemple : les méthodes d'améliorations textuelles, page 132) ou encore l'emploi de différents supports (exemple : désignation par caméra déportée, page 145) nous montrent des voies de recherches prometteuses.

Perspectives

Comme nous l'avons soutenu et illustré, les possibilités d'agir sur les communications médiatisées par un monde virtuel sont nombreuses. Les perspectives peuvent s'inscrire suivant différents axes de recherche.

Le premier axe est directement lié à un manque de nos travaux : **la validation**. L'illustration de nos améliorations par un scénario (Chapitre 6) nous montre les apports possibles de l'agent. Toutefois, il nous faut nous garder de généraliser la portée de ces résultats avant une validation.

Le second axe concerne les différentes étapes de traitement d'un acte de communication.

- traitement en entrée : notre **reconnaissance des messages textuels** est simple et concerne des phrases courtes émises par l'émetteur. Les phrases longues ou complexes ne sont pas traitées. Cette amélioration permettrait ainsi l'extension des possibilités d'intervention de l'agent.
- traitement de la communication : elle passe par la mise en œuvre de **nouvelles méthodes de traitement** des informations disponibles au niveau l'environnement virtuel. Elle concerne aussi bien une étude au niveau de l'architecture du support que de la structure du format d'enregistrement des informations. Nous pouvons aussi ajouter les possibilités d'améliorer les méthodes existantes. Par exemple, nous pouvons améliorer la méthode de classification des éléments d'une scène virtuelle. Ainsi, avec ce classement, nous pouvons générer des informations plus pertinentes sur le positionnement géographique d'un objet.
- traitement en sortie : le **choix des méthodes** utilisées pour obtenir des informations, est effectué par le récepteur. Il apparaît que le choix de présélection de l'agent (par exemple dans le cas des améliorations textuelles) peut être effectuée en fonction de comportements antérieurs du récepteur et d'autres paramètres virtuels comme sa position vis-à-vis des éléments de la scène. Nous pourrions ainsi fournir des informations plus appropriées pour le récepteur.

Un troisième axe concerne l'exploitation de l'architecture de l'application et de la scène virtuelle pour l'élaboration d'un agent, comme dans les travaux de J.C. Martin[44]. Il reste à rendre **l'agent plus adaptable** aux différents environnements virtuels. C'est dans ces conditions que nous pourrions envisager l'application de notre modèle d'agent à d'autres mondes virtuels.

Enfin, dans un contexte plus théorique, il semble possible d'**élargir les niveaux de communication traités**. Pour notre part, nous nous sommes consacrés au traitement des communications interpersonnelles. Il serait intéressant de pouvoir traiter tous les niveaux de communication décrits par la pyramide de McQuail (figure 2.1 page 25). L'élargissement des améliorations aux autres niveaux de communication nous semble se situer dans le prolongement de notre approche.

8 Annexes

Sommaire

8.1	Autorisations Images	173
8.2	Analyse du mode de «picking»	177
8.3	Positionnement automatique	180
8.4	Transmission d'images à travers un réseau . . .	182

Cette partie du manuscrit regroupe les différentes annexes citées tout au long de nos écrits. Les unes permettent seulement à étayer un argumentaire. Alors que d'autres permettent d'argumenter certains choix ou affirmations.

8.1 Autorisations Images

Dans cette section, nous montrons les différentes autorisations existantes et obtenues pour les images. Toutes ces informations ont été obtenues par courriel ou recherche sur internet. Veuillez-me contacter en cas de modification de licence.

Licence Publique Générale

Toute image provenant de programmes sous GPL qui ne contient aucune information personnelle est normalement sous licence publique générale. Toutes les images sont alors en diffusion libre sous licence GPL. La modification est autorisée s'il reste sous licence GPL.

<i>Figure 3.17 page 69</i>	Open Croquet 3D
<i>Figure 3.25 page 82</i>	BZFlag
<i>Figure 3.10 page 57</i>	X-Chat
<i>Figure 3.9 page 56</i>	Dive
<i>Figure 3.14 page 63</i>	Open Croquet 3D

TAB. 8.1 – Les images sous licence GPL

License WikiCommons et libre Diffusion

Les images de cette section possèdent des licences de libre diffusion à but non-commercial. De manière général, la diffusion est autorisée. N'ayant pas eu besoin de les modifier, les autorisations sont à demander en cas de modification pour une diffusion.

<i>Figure 3.1 page 47</i>	http://upload.wikimedia.org/wikipedia/en/e/e4/WoW-Lakeshire.png
<i>Figure 3.2 page 47</i>	http://www.evl.uic.edu/pape/CAVE/oldCAVE/caveback_small.jpg
<i>Figure 3.3 page 48</i>	http://lsiit.u-strasbg.fr/sites/img/dogmerv.jpg
<i>Figure 3.4 page 49</i>	http://lsiit.u-strasbg.fr/sites/img/phantom_desktop.jpg
<i>Figure 3.6 page 51</i>	http://www.dgp.utoronto.ca/people/modjeska/Pubs/iv2000_genetics.pdf
<i>Figure 3.7 page 52</i>	http://www.virtual.gmu.edu/pdf/aied.pdf
<i>Figure 3.8 page 52</i>	http://portal.acm.org/citation.cfm?id=351006.351010
<i>Figure 3.21 page 78</i>	http://slhistory.org/images/6/69/Sheridan_today.jpg
<i>Figure 3.22 page 79</i>	http://upload.wikimedia.org/wikipedia/en/0/04/TheSims.jpg
<i>Figure 3.25 page 82</i>	http://upload.wikimedia.org/wikipedia/en/thumb/f/fc/ATM000251.jpg/800px-ATM000251.jpg
<i>Figure 3.19 page 74</i>	http://www.informatics.bangor.ac.uk/~nigel/eg2003-mp/6/soler1.jpg
<i>Figure 3.20 page 76</i>	http://www.limsi.fr/~jps/demos/picts/LEA.JPG

TAB. 8.2 – Les images sous Licence de diffusion libre

Autorisations des auteurs

Ces images possèdent un copyright et des mails ont été envoyés aux auteurs de ces images pour obtenir l'autorisation d'apparaître dans ce document. Ces images ont été autorisées par les auteurs. La modification ou même la diffusion n'est pas autorisée.

<i>Figure 3.4.7 page 64</i>	Autorisation des auteurs des articles de MERL
<i>Figure 3.4.10 page 71</i>	Autorisation des responsables de : //www.haboo.fr
<i>Figure 3.1 page 47</i>	Autorisation d'un responsable de : //www.karobas.fr , société Karobas - Saint Bernard , Parc des Moulins 2 - Porte 19, 2791 Chemin de St Bernard 06220 VALLAURIS
<i>Figure 3.25 page 82</i>	Autorisation d'un concepteur de du site : ://www.kochonland.fr , tous droits réservés, ©L.Sarrazin, O. Coudre, Explore, 3nko, INPI 2003
<i>Figure 3.25 page 82</i>	Autorisation des auteurs, diffusion autorisé car non-commercial
<i>Figure 3.5 page 50</i>	Autorisation d'un développeur, tous droits réservés.
<i>Figure 3.23 page 80</i>	Autorisation de la mairie de Village3D
<i>Figure 3.24 page 81</i>	Autorisation de VR4all contre une copie du document.
<i>Figure 3.4.8 page 67</i>	Autorisation d'un responsable d'activeworld.

TAB. 8.3 – Les images avec autorisations spécifiques

Divers

Les images et schémas non cités sont soit dans le domaine public ou sont issues de nos travaux.

8.2 Analyse du mode de «picking»

Dans cette partie, nous allons expliquer une méthode utilisée à deux reprises pour obtenir des informations sur l'occultation des objets.

L'élaboration de cette méthode peut s'effectuer suivant deux approches possibles. La première approche est le passage par la méthode du lancé de rayon. Elle consiste à lancer un rayon depuis un point de vue et à analyser les objets traversés par ce rayon. La seconde approche est la méthode du picking qui consiste à récupérer l'ensemble des objets dessinés à l'écran.

La méthode du lancé de rayon fonctionne parfaitement, néanmoins, nous ne profitons pas de la puissance des cartes graphiques car chaque test est effectué par le *CPU*. Le calcul est fortement ralenti lorsque la scène présente un nombre important d'objets.

Un des avantages de cette méthode est que nous pouvons gérer des intersections en profondeur afin d'obtenir des informations détaillées. Mais cette méthode nécessite une implémentation plutôt lourde alors qu'il existe une méthode plus simple. La seconde méthode existe sous le nom de « picking » et permet de dépasser les problèmes que nous pouvons rencontrer dans le lancé de rayon (par exemple : les difficulté pour le lancé de rayon de détecter les surfaces planes perpendiculaires). Cette seconde méthode a donc été choisi pour calculer le pourcentage d'occultation.

OpenGL prévoit deux modalités pour le picking selon le type d'information que l'on souhaite récupérer : le mode `GL_SELECT`, qui permet de savoir quel objet ou quelle face est visible, et le mode `GL_FEEDBACK`, qui permet d'obtenir toutes les informations disponibles sur la scène, coordonnées géométriques dans le repère écran, couleur, coordonnées textures (u,v)... On peut par exemple en tirer les coordonnées du point cliqué à la souris à la surface d'un objet. On active ce mode par `glRenderMode(mode)` (le mode de tracé normal avec affichage correspondant au mode `GL_RENDER`).

Dans les deux cas il faut fournir un buffer où seront stockées les informations retournées par le geometric engine, avec selon le mode `glSelectBuffer(size, buff)` ou `glFeedbackBuffer(size, GL_3D_COLOR_TEXTURE, buff)`. On accède à ce buffer après être sorti du mode picking par l'appel de la méthode : `nbinfo = glRenderMode(GL_RENDER)`.

Le mode `GL_SELECT`

Dans le mode `GL_SELECT`, on regroupe les primitives qui constituent une unité logique en insérant des déclarations `glPushName(i)` entre leurs tracés, ce qui revient à donner un identifiant à ces unités. Le geometric en-

gine – portion du graphics engine assurant les transformations géométriques (changements de repère, projection. . .). – inscrit dans le buffer l’identifiant de toute unité dont au moins une primitive se projette dans la zone de l’écran, c’est ainsi que l’on sait quel objet est dessiné à l’écran.

Le mode `GL_FEEDBACK`

Dans le mode `GL_FEEDBACK`, chaque primitive apparaissant potentiellement à l’écran va générer une entrée dans le buffer, consistant en un identifiant du type de primitive, `GL_POLYGON_TOKEN` pour un polygone, suivi du nombre de sommets, puis pour chaque sommet, de tous les attributs demandé, soit pour `GL_3D_COLOR_TEXTURE` les coordonnées x,y,z dans le repère écran, la couleur apparente `RGBA`, et les coordonnées de texture u,v,s,t (dans le cas général, il y a 4 coordonnées texture). Comme la phase de clipping redécoupe les facettes qui sortent de l’écran, dans le cas d’une fenêtre de quelques pixels de large positionnée sous la souris on obtient ainsi les coordonnées, l’illumination. . ., sur le lieu de la surface situé directement sous la souris. NB : on peut remonter aux coordonnées dans le repère monde via la fonction `gluUnproject()`. (Mais cette fonction est coûteuse si l’on veut retrouver plusieurs points, car elle inverse la matrice de transformation à chaque appel. Mieux vaut alors faire soit même l’inversion de matrice pour appliquer la même transformation réciproque à tous les points.)

Discussion sur les deux modes de picking

On peut aussi noter que l’on obtient à l’aide de ce mode un descriptif de ce qui est « visible », et en aucun cas la désignation des objets comme dans le mode `GL_SELECT`.

Il reste possible de retrouver cette fonctionnalité en insérant des appels à `glPassThrough(f)` entre les primitives constituant une unité, dans le même esprit que `glPushName(i)`. On récupère alors dans le buffer l’identifiant de primitive `GL_PASS_THROUGH_TOKEN` suivi de la valeur `f` passée en argument, ce qui permet de reconnaître l’unité à laquelle appartiennent les entrées du buffer. Mais toutes ces fonctionnalités sont-elles vraiment utiles pour notre amélioration ?

En analysant en détail le mode `GL_FEEDBACK`, nous observons qu’il apporte des informations supplémentaires. Ces informations riches et complexes ne sont pas exploitées par notre agent. Dans le cadre de notre recherche, il serait très intéressant de calculer la forme perçue mais ceci demande un calcul trop important que nous n’avons pas choisi d’implémenter.

Les informations de couleurs et de textures perçus peuvent être une

source d'informations pour l'agent. il pourrait les utiliser pour comprendre ce qui est perçu dans le champ visuel des locuteurs. Bref, ces informations permettent une bonne précision sur les couleurs qui nous est peu utile. Mais cette précision est tronquée par les fichiers de gestion de couleurs (icc, icm. . .) qui sont parfois mal gérés par les cartes et les drivers, on peut ajouter à cela les problèmes de couleurs liées à l'écran lui même. En bref, cette précision devient peu utilisable et il en va de même pour les informations de textures.

En conclusion, le mode `GL_SELECT` semble être le plus adapté à notre recherche et fournit des informations suffisantes pour notre emploi.

8.3 Positionnement automatique

Dans cette partie, nous allons expliquer le calcul du positionnement automatique. Nous expliquons ici le but de notre approche et l'implémentation choisie.

Nos méthodes de visualisation tiennent compte d'un placement automatique de la caméra de visualisation. Certains environnements virtuels possèdent des informations de positionnement nommées. Cette méthode a l'avantage de donner clairement un point de vue sur un objet virtuel possédant des informations de points de vues. Mais dans un monde où les objets se créent et se détruisent constamment, cette méthode montre une certaine difficulté. Il nous est difficile de créer et de détruire constamment des points de vues sur les objets. Dès lors, nous avons recherché une méthode pour calculer automatiquement une position de visualisation pour un objet virtuel.

Explications

L'objectif est d'obtenir une méthode permettant de fournir pour un objet quelconque de la scène, des coordonnées «optimales» de visualisation. Nous allons expliquer les choix techniques et la méthode qui nous a permis d'atteindre notre objectif.

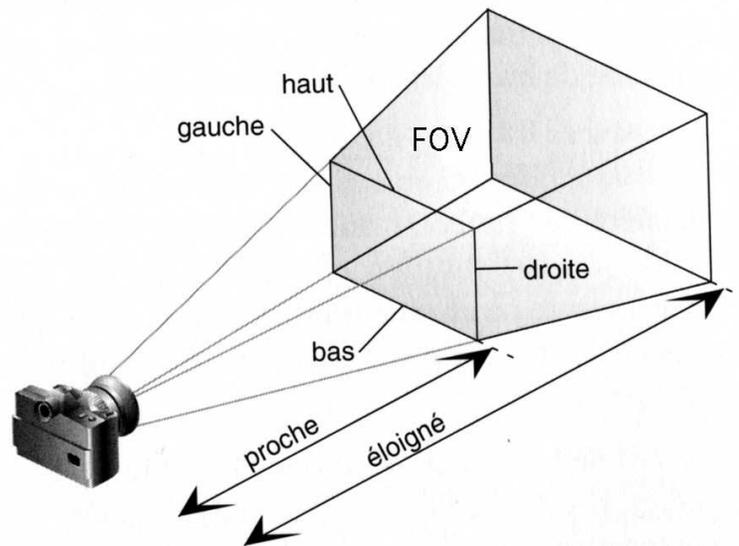


FIG. 8.1 – *Le champ de vision sous OpenGL*

Sous OpenGL, lorsque nous visualisons une scène, nous pouvons décrire

le champ de vision comme une pyramide à base rectangulaire tronqué (cf. page 116 livre OpenGL [109] cf. figure 8.1). Il est à noter que nous pouvons connaître les paramètres (taille) de ce champ de vision. Nous savons aussi que nous pouvons obtenir les coordonnées de la «boîte de collision» de n'importe quel objet du monde virtuel.

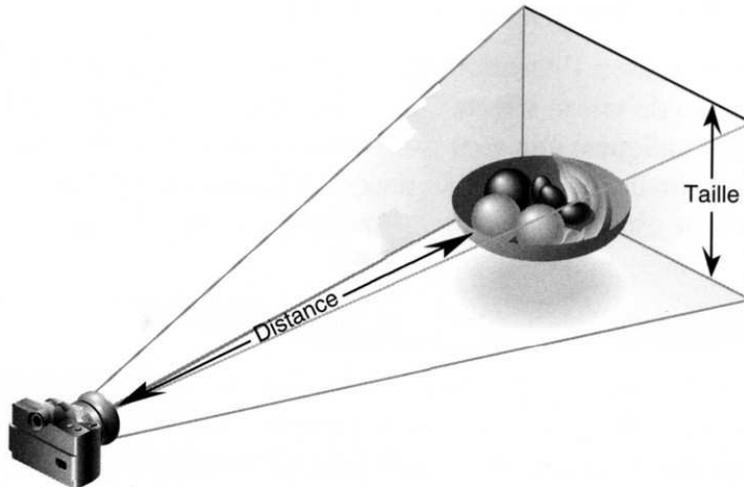


FIG. 8.2 – *La recherche de positionnement*

Nous souhaitons obtenir un positionnement automatique pour obtenir une vision sur un objet recherché (exemple cf. figure 8.2). Le calcul de la position automatique s'effectue à l'aide des coordonnées de l'objet et des propriétés du champ de vision. La méthode est de calculer des coordonnées et des orientations du champ de vision afin d'adapter la visualisation à la boîte de collision. Notre méthode va rechercher des coordonnées de visualisation pour que la boîte de collision de l'objet remplisse une partie de la pyramide tronquée. Le résultat est un ensemble de coordonnées équidistants du centre de l'objet étudié. De cet ensemble, nous retirons les coordonnées non pertinentes : les coordonnées trop hautes, les coordonnées trop basses et les orientations peu probables. De cet nous tirons quatre groupes de coordonnées et d'orientations qui répondent à des critères de logiques : ils sont à des hauteurs de similaires mais ont des orientations très différentes.

Pour chaque groupe de résultat, nous simulons la scène en arrière-plan en calculant le pourcentage d'occultation de l'objet afin de choisir la première méthode dépassant un pourcentage d'occultation pertinent.

8.4 Transmission d'images à travers un réseau

Dans nos améliorations, nous avons souhaité pouvoir améliorer le support de communication notamment en offrant la possibilité d'adjointre une image au message textuel. Ce souhait a soulevé un problème réseau que nous avons essayé de dépasser et de solutionner.

La transmission d'une image de l'émetteur au récepteur est une approche difficile à élaborer. Regardons quelques informations techniques connus. Actuellement, les résolutions d'écran utilisées varient entre $1024 * 768$ et $1600 * 1200$ avec pour la plupart à 16 bits de couleurs[110]. L'image d'une scène posséderait une taille de 1,5 à 3 Mo sans compression. Dans la pratique avec la compression avec des pertes (ex : format «jpeg»), nous avons des fichiers images de 300Ko à 700Ko. Par la suite, il nous fallait envoyer cette image au récepteur.

De nombreux problèmes sont apparus dont celui de la sécurité et des autorisations au niveau du récepteur. La mise en place de cette méthode suivant cette approche s'est trouvé dans une impasse. Nous ne pouvions transmettre aisément des images à travers le réseau. VREng ne permettait pas une mise en place simple. Dès lors, nous nous sommes orientés vers une autre approche.

Nous sommes au sein d'un monde virtuel, ce contexte virtuel dans les systèmes multi-utilisateurs offre un contexte commun à tous les utilisateurs. Nous avons alors choisi de transmettre des coordonnées au lieu des images. La transmission de ces coordonnées s'effectue aisément en même instant que les messages textuels. Le récepteur reçoit les coordonnées, calcule l'image à générer afin d'obtenir l'image que souhaite transmettre l'émetteur.

Lorsque l'émetteur transmet des coordonnées, le récepteur utilise une méthode de «off-screen rendering» pour simuler et générer l'image dont voici une explication technique.

La méthode « off-screen rendering »

La méthode est assez simple et connue par les spécialistes 3D, elle permet d'effectuer des simulations non visibles pour diverses applications. Cette méthode est communément utilisée pour un affichage en « double-buffer »¹

OpenGL et DirectX permettent un emploi personnel de ces buffers. ils

¹Le double buffer est une technique ayant pour but d'amélioration la qualité et la vitesse d'affichage lors de la réalisation animations où les demandes de réaffichage sont importantes.

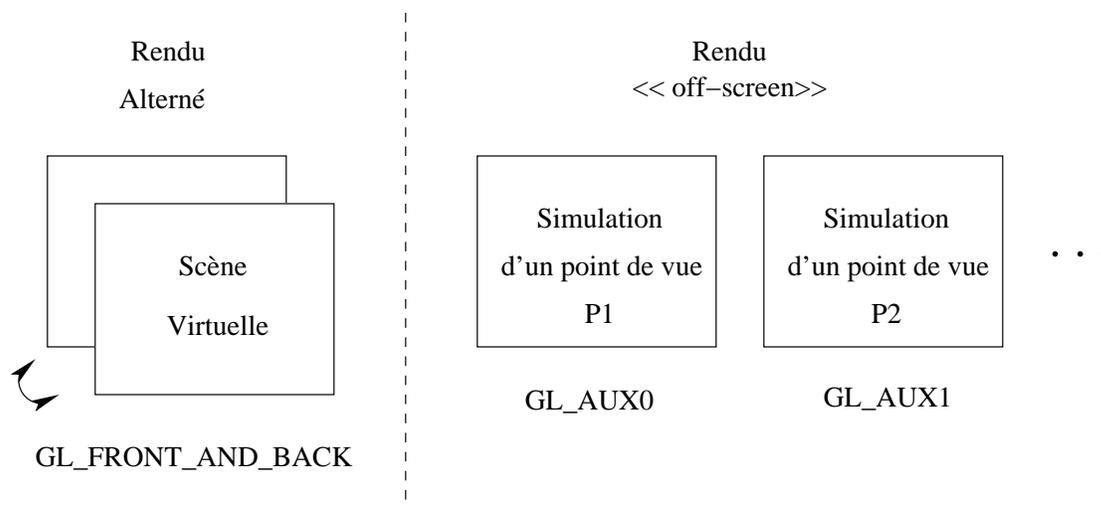


FIG. 8.3 – Les tampons d’affichage sous OpenGL

existent non seulement les deux buffers d’affichage classique pour le double-buffer mais aussi des buffers pour la vision stéréo et des buffers auxiliaires pour les calculs spécifiques. À l’aide des coordonnées reçues, nous simulons la scène vue depuis ces coordonnées dans l’un des buffers auxiliaires « GL_AUXn » (cf. Figure 8.3). L’image n’apparaît pas à l’écran mais pourtant elle existe. Il nous reste alors à enregistrer la vision ainsi calculée et à l’enregistrer dans un fichier afin de pouvoir de visualiser. Nous obtenons ainsi l’adjonction d’un image au message textuel tel que souhaitée par l’émetteur.

Bibliographie

- [1] Michael ABRASH : *Graphics Programming Black Book*. Dr. Dobb's Journal, 2006. programmation avec ModeX de différents concepts classique en 3D. **Résumé** : Livre regroupant de nombreux concepts 3D connus mais adaptés à des problèmes très spécifiques. Il permet de comprendre les limites de certaines méthodes de calculs utilisant le graphisme et les spécificités d'une simulation 3D.
- [2] Jean-Claude ABRIC : *Psychologie de la communication : Théories et méthodes*. Armand Colin Editions, 2^e édition, 2003. ISBN : 2200265565. **Résumé** : Ce livre fournit des explications sur différents autres modèles de communication. Il explique les théories mises en jeu dans ces modèles.
- [3] Bernd AMANN et Philippe RIGAUX : Transformation de documents xml. In O'REILLY, éditeur : *Comprendre XSLT*, pages 63–175, 276–300, 359–409. O'Reilly, 2002. XML, XSLT, DTD-Xpath, XML dynamique XSP et XSQL, **Résumé** : Ce livre généraliste permet de comprendre les apports possibles de XML. Il fournit un exemple complet sur les transformations auxquelles nous pouvons procéder sans grandes ressources.
- [4] David B. ANDERSON, John W. BARRUS, John H. HOWARD, Charles RICH, Chia SHEN et Richard C. WATERS : Building multi-user interactive multimedia environments at merl. In *IEEE Multimedia*, volume 2, 4, pages 77–82, Décembre 1995.
- [5] Denis ARNAUD, Fabrice BELLAR, Stéphane BELMON, Samuel ORZAN et Lionel ULMER : Vreng : Analyse globale. Rapport technique, ENST, 1998. **Résumé** : VREng est une plate-forme de développement provenant des travaux des étudiants de l'ENST et à laquelle Philippe Dax est à l'origine. L'objectif de cette plate-forme est multiple, il s'agit de fournir un univers permettant le développement simple d'outils pour les mondes virtuels.
- [6] Jean-Christophe BERNADAC et François KNAB : Etude de cas. In Inc. EYROLLES, éditeur : *Construire une application XML*. CosmoBay L'esprit Système Internet., 1999. **Résumé** : Permet la compréhension de

- XML à travers divers exemples. Livre instructif sur XML et des outils qui l'entourent.
- [7] Olivier BERSOT, Pierre-Olivier EL GUEDJ, Christophe GODÉREAUX et Pierre NUGUES : A conversational agent to help navigation and collaboration in virtual worlds. *Virtual Reality*, 3(1):71–82, 1998.
- [8] Yann BERTACCHINI : territoire physique/territoire virtuel quelle cohabitation ?, 2003. Le Groupe GOING (Groupe d'observation et d'Investigation des Nouvelles Gouvernances).
- [9] Franck BIOCCA et Mark R. LEVY : *Communication in the age of Virtual Reality*. Lawrence Erlbaum Associates, Inc., 1996. ISBN : 0805815503. **Résumé** : Livre généraliste sur la réalité virtuelle, de nombreuses recherches sont expliquées notamment sur les différents problèmes apparus sur les cyber-instruments. Il fournit des axes de recherche et offre un aperçu général de la réalité virtuelle.
- [10] Bruno BOSSARD, Thomas CONVARD, Annelies BRAFFORT, Damien TOURAINE, Patrick BOURDOT et Michèle JARDINO : A generic gesture recognition system for virtual reality. In *RFIA 2004*, Toulouse, France, 2004. RFIA 2004, 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle. <http://citeseer.csail.mit.edu/braffort01research.html>, **Résumé** : Article concernant la reconnaissance de gestes. On y retrouve des explications sur l'utilisation de gants cybergloves pour reconnaître certains avec les limites de ce type de reconnaissance.
- [11] Patrick BOURDOT et Jean LORISSON : Rapport final de présentation pour une plateforme de réalité virtuelle en France. Rapport technique, CNRS, Octobre 2004. <http://www.perfrv.org>.
- [12] Jean-Yves BOYER : Influence de la structure textuelle et de la redondance sur la lisibilité. *Canadian Journal of Education / Revue canadienne de l'éducation*, pages 16–36, Décembre 1990. Vol. 15, No. 1, doi :10.2307/1495415.
- [13] I. BRETAN et J. KARLGREN : Worlds without words. In *Proceedings of ERCIM Workshop on Virtual Reality, SICS, Stockholm*, 1994. <http://citeseer.ist.psu.edu/bretan99worlds.html>.
- [14] Ivan BRETAN : Natural language in model world interfaces. Rapport technique R95-03, Department of Computer and Systems Sciences, The royal Institute of Technology and Stockholm University, 1995. <http://citeseer.ist.psu.edu/bretan95natural.html>, **Résumé** : Dans ce manuscrit, une partie des travaux concerne la description des champs d'actions pour les modalités. Il décrit notamment la faible utilité de certains supports suivant le contexte visuel. Une description intéressante de la force et des faiblesses du langage naturel.

- [15] Stéphanie BUISINE, Abrilian SARKIS et Jean-Claude MARTIN : Evaluation of multimodal behaviour of embodied agents. *In Kluwer Academic Publishers*, pages 217–238, 2004.
- [16] Judée K. BURGOON : *A Communication model of personnel space violations : explication and an initial Test*, pages 129–142. Ed. D., University of Florida, Gainesville, Florida 32611, Janvier 1978. Volume 4, Issue 2, **Résumé** : La distance de communication est un facteur qui apporte certains types d'informations qui auraient pu nous permettre de découvrir de nouvelles méthodes d'analyse des conversations en tenant compte des distances entre les utilisateurs. Ce livre présente des modèles. il fournit des explications sur la distance de communication et des orientations de recherches possibles.
- [17] Jean CAELEN et Anne XUEREBO : *Interaction et Pragmatique : jeux de dialogue et de langage*. Hermes Sciences, Lavoisier, Collection science informatique et SHS, 2007. **Résumé** : Ce livre traite de l'interaction Homme-machine (plus précisément du dialogue humain) de manière théorique, méthodologique et pratique. L'ouvrage met l'accent sur l'interprétation. Il offre un parcours théorique, les questions de multimodalité et d'évaluation sont abordées, donnant de nouvelles perspectives de recherche et des moyens de diagnostiquer les performances des systèmes.
- [18] Christophe CHAILLOU, A. LUCIAN, A. HABIBI et J.P. JESSEL : State of art : image synthesis and virtual reality. *In Virtual Reality and prototyping, Laval Virtual*, France, 1999. **Résumé** : Cet article décrit des axes de recherches en réalité virtuelle notamment les apports et les limites de la réalité virtuelle dans différents aspects tels que : l'accès à un contact sensoriel avec un élément du monde inaccessible, dû à des facteurs temporels (l'objet n'existe plus) ou encore des facteurs géographiques (l'objet est éloigné ou est de taille trop différente, par exemples des molécules ou un univers.) L'article se décline suivant trois axes : modélisation/simulation, création de supports, et étude des interactions entre utilisateur et environnement.
- [19] Emmanuel CHAILLOUX, Pascal MANOURY et Bruno PAGANO : *Objective CAML*. O'Reilly, Février 2004. ISBN : 2841771210. **Résumé** : L'implémentation de notre agent s'est faite en Ocaml, ce livre complète le livre en ligne de l'INRIA qui explique les approches facilitées par ce langage et donne une base stable pour la programmation d'un analyseur syntaxique.
- [20] M. CHOI et J. CREMER : Interactive manipulation of articulated objects with geometry awareness. *In Proceedings of IEEE International Conference on Robotics and Automation.*, Mai 1999. <http://citeseer.ist.psu.edu/choi99interactive.html>, **Résumé** : Présentation de techniques de manipulation des objets en 3D dans des environnements ré-

- partis. Les auteurs présentent des méthodes permettant de manipuler des objets en évitant les interpénétrations. L'article se concentre sur les détections de collisions et les interpénétrations.
- [21] Web 3D CONSORTIUM : X3d overview, 2004. X3D : eXtended 3D, **Résumé** : Site d'apprentissage du langage X3D. Il regroupe des exemples et des guides pour apprendre, comprendre et utiliser X3D. Toutes les connaissances et les normes y sont regroupées.
- [22] A. CORRADINI et P. COHEN : On the relationships among speech, gestures, and object manipulation in virtual environments : Initial evidence. *In Proceedings of Int'l CLASS Work. on Natural, Intell. and Effective Interaction in Multimodal Dialogue Systems*, 2002. <http://citeseer.csail.mit.edu/corradini02relationships.html>, **Résumé** : c'est une présentation d'un magicien d'Oz dont le but a été de comprendre les gestes et les paroles lors de jeux vidéos. Les résultats présentés sous la forme d'un tableau, présente les différents emplois suivant les environnements utilisés. L'analyse fournit le pourcentage de temps passé sur les modalités.
- [23] Bob COYNE et Richard SPROAT : Wordseye : An automatic text-to-scene conversion system. *In SIGGRAPH*, éditeur : *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 487–496, Los Angeles, CA, USA, Février 2002. **Résumé** : le langage naturel est utilisé comme un outil de description de scènes et d'images mentales. C'est à partir de cette idée que le logiciel WordsEye a été créé. Ce système converti automatiquement un texte en une représentation 3D. Le logiciel fait un lien entre le texte qui lui est rentré et une base de données de modèles 3D. En effectuant une analyse complexe des phrases. Ils utilisent les propriétés des objets pour évaluer l'agencement des objets les uns en fonction des autres. L'objectif final est d'obtenir une scène proche de la scène imaginée et décrite. Cet écrit contient une étude importante sur la forme des objets et leurs propriétés d'interactions.
- [24] Jean-François DORTIER : *Le dictionnaire des sciences humaines*, volume 9 de 3. Sciences humaines, 2004. **Résumé** : C'est un livre généraliste sur la communication avec une approche simple permettant de comprendre au mieux les différents concepts et approches de la communication. Nous avons repris de nombreuses citations de ce livre ainsi que quelques exemples. Ce livre est un point de départ intéressant pour aborder la communication.
- [25] Jim DURBIN, J. Edward Swan II, Brad COLBERT, John CROWE, Rob KING, Tony KING, Christopher SCANNELL, Zachary WARTELL et Terry WELSH : Battlefield visualization on the responsive workbench. *In IEEE Visualization*, pages 463–466, 1998. <http://citeseer.ist.psu.edu/durbin98battlefield.html>, **Résumé** : Cet article présente

- un système de visualisation dénommé «Dragon». Il a été élaboré sur un WorkBench. Ce système a été utilisé avec succès lors d'un exercice militaire. Il permet de manipuler et de visualiser rapidement de manière efficace un champ de bataille. Le système est multimodal, il utilise un pinchglove, une reconnaissance vocale et un joystick. Une grande partie présente des paradigmes pour navigation et la manipulation d'objets.
- [26] P. EKMAN, W.V. FRIESEN, M. O'SULLIVAN et K. SCHERER : Relative importance of face, body, and speech in judgments of personality and affect. *Journal of Personality and Social Psychology*, 38:270–277, 1980.
- [27] Stephanie EVERETT, Kenneth WAUCHOPE et Manuel PEREZ : A natural language interface for virtual reality systems. *In Virtual Reality Systems*, <http://www.aic.nrl.navy.mil/severett/>, 1996. <http://cite-seer.ist.psu.edu/everett96natural.htm>, **Résumé** : Cet article présente le robot «Robart III». Ce robot possède un système d'analyse du langage naturel dont le but est de comprendre les requêtes en langage naturel et d'effectuer l'action recherchée. C'est un robot utilisé pour l'exploration de zones dangereuses. L'exemple utilisé est celui d'un groupe de soldats bloqués devant une porte et ne pouvant pas avancer. Le robot reçoit l'ordre de : «avance vers la porte et regarde autour». Le robot avance vers la porte et effectue l'action.
- [28] W.N. Street F. MENCZER et A.E. MONGE : Adaptive assistants for customized e-shopping. *In IEEE Intelligent Systems and Their Applications*, pages 12–19, University of Iowa, Ames, IA, USA, Décembre 2002. Volume : 17, Issue : 6, ISSN : 1541-1672, INSPEC Accession Number : 7485120, Digital Object Identifier : 10.1109/MIS.2002.1134357, Posted online : 2005-04-04 09 :38 :05.0, **Résumé** : Cet article présente un modèle d'agent intelligent pour les achats en ligne. Il utilise un système de préférences entre les sites et l'utilisateur du système. Il a pour objectif de rechercher et de proposer des articles qui peuvent intéresser l'utilisateur.
- [29] Jacques FERBER : *Les systèmes multiagents : Vers une intelligence collective*. Inter-Editions, 1995. **Résumé** : Livre général sur les systèmes multi-agents (SMA). Il propose de nombreuses approches permettant de comprendre les différents concepts et des applications possibles. Un livre de référence pour ces systèmes.
- [30] Jaques FERBER : Les systèmes multi-agents : un aperçu général. *In Technique et Science Informatique*, volume 8, pages 979–1012, 1997.
- [31] Frans FLIPPO, Allen KREBS et Ivan MARSIC : A framework for rapid development of multimodal interfaces. *In Proceedings of the 5th international conference on Multimodal interfaces*, volume 5, pages 109–116. Association for Computing Machinery, ACM Special Interest Group on Computer-Human Interaction, ACM Press, New York, NY USA, 2003.

ISBN : 1-58113-621-8, **Résumé** : Malgré l'accessibilité d'interfaces multimodales, il n'existe pas ou peu d'applications commerciales utilisant la multimodalité. L'une des causes est le manque de supports permettant un développement rapide sur ce type d'interfaces. Cet article décrit un framework permettant l'implémentation rapide d'interfaces multimodales avec peu de ressources.

- [32] Andrew U. FRANK : Qualitative spatial reasoning : Cardinal directions as an example. *International Journal of Geographical Information Science*, 10(3):269–290, 1996.
- [33] Stan FRANKLIN et Art GRAESSER : Modeling cognition with software agents. *In Institute for Intelligent Systems*. The University of Memphis, Memphis TN 38152, US, 2001. **Résumé** : Cet article présente l'emploi d'agents autonomes comme modèles cognitifs. Ils génèrent des hypothèses sur le modèle cognitif Humain. L'objectif est de produire des aspects typiquement humain afin de les évaluer.
- [34] Mike FRASER, Tony GLOVER, Ivan VAGHI, Steve BENFORD, Chris GREENHALGH, Jon HINDMARSH et Christian HEATH : Revealing the realities of collaborative virtual reality. *In Proceedings of the third international conference on Collaborative virtual environments*, pages 29–37, San Francisco, California, United States, 2000. Association for Computing Machinery Press, New York, USA. ISBN :1-58113-303-0, <http://portal.acm.org/citation.cfm?id=351010>, **Résumé** : On tente ici de montrer les différences entre le monde réel et le monde virtuel. Ceci passe par une explication des limites de chacun et surtout des limites techniques de la réalité virtuelle. Ils nous présentent quelques illustrations pour nous révéler les limites.
- [35] E. FRECON, H. J-ARO et S. STENIUS : Dive - The Distributed Interactive Virtual Environment - dive files description for dive version, 1997. <http://citeseer.ist.psu.edu/avatare97dive.html>.
- [36] Emmanuel FRECON et Marten STENIUS : Dive : A scaleable network architecture for distributed virtual environments. *In the British Computer Society The Institute of PHYSICS et the Institution of ELECTRICAL ENGINEERS*, éditeurs : *Distributed Systems Engineering*, pages 91–100, Swedish Institute of Computer Science Box 1263 SE-164 28 Kista, Septembre 1998. **Résumé** : Dive une plate-forme de réalité virtuelle multi-utilisateurs, où les participants peuvent naviguer dans un univers 3D. Ils peuvent se promener, rencontrer et interagir sur les différents éléments du monde. Les recherches et applications de Dive sont nombreuses, citons notamment les interactions multi-modales , la visualisation de champ de bataille, le contrôle de robot dans le monde réel.

-
- [37] Emmanuel FRÉCON, Olov STÅHL et Anders WALLBERG : The dive client interface - reference document. <http://cite-seer.ist.psu.edu/300695.html>.
- [38] Olivier GÉNEVAUX, Arash HABIBI et Jean-Michel DISCHLER : Simulating fluid-solid interaction. In *Graphics Interface*, pages 31–38. CIPS, Canadian Human-Computer Communication Society, A.K. Peters, Juin 2003. ISBN 1-56881-207-8, ISSN 0713-5424.
- [39] George GERBNER : Toward a general model of communication. *AV Communication Review* 4, 4:171–199, Juillet 1956.
- [40] George GERBNER, Ole R. HOLSTI, Klaus KRIPPENDORFF, William J. PAISLEY et Phillip STONE : *The Analysis of Communications Content : Developments in Scientific Theories and Computer Techniques*. John Wiley and Sons, New York, Etats-Unis, 1969.
- [41] Alain GRUMBACH : *Cognition Artificielle*. Editions Addison-Wesley France, 1994. Du reflexe ... à la réflexion, ISBN : 2879080592.
- [42] Alain GRUMBACH : *Cognition Virtuelle : Réflexion sur le virtuel, ses implications cognitives et ses réalisations artistiques*. GET ENST Paris, 2003.
- [43] INRIA : Mondes réel - mondes virtuels, proposition d'axe stratégique. Rapport technique, INRIA, 2003. **Résumé** : Rapport technique détaillant les motivations de la recherche sur les langages de description. nous y retrouvons des explications sur différents axes de recherches notamment la représentation du monde réel et construction de mondes virtuels.
- [44] Karl Leguern JEAN-CLAUDE MARTIN, Jean-Paul Sansonnet : Intégration de l'agent lea du limsi à l'architexture dialogique daft. In *LIMSI/CNRS*, 2005. **Résumé** : Présentation d'une architecture dialogique médiateur pour une agent d'assistance conversationnel animé. Présentation du projet ACA appelé LEA.
- [45] Michel JOSIEN : *Techniques de communication interpersonnelle*. eds d'Organisation, Janvier 2004. Collection Travail A Grande Efficacite, ISBN : 270813051X.
- [46] Ed KAISER, Alex OLWAL, David MCGEE, Hrvoje BENKO, Andrea CORRADINI, Xiaoguang LI, Phil COHEN et Steven FEINER : Mutual disambiguation of 3D multimodal interaction in augmented and virtual reality. In *ICMI'03*, ACM : 1-58113-621-8/03/0011, Vancouver, British Columbia, Canada, Novembre 2003. Oregon Health and Science University, ICMI'03.
- [47] J. KARLGREN, I. BRETAN, N. FROST et L. JONSSON : Interaction models, reference, and interactivity in speech interfaces to virtual environments. In *Proceedings of Eurographics'95*, 1995. <http://cite-seer.ist.psu.edu/300695.html>.
-

- seer.ist.psu.edu/karlgren95interaction.html, **Résumé** : Après une présentation du problème de la désignation d'un objet, on présente les intérêts d'employer la modalité vocale dans les environnements virtuels. Suivant les situations, cette modalité est un ajout qui renforce l'impression d'immersion. Différents exemples des champs d'utilisation des modalités sont présentés ainsi que des apports de l'emploi synergique.
- [48] Evelyne KLINGER : *Apports de la réalité virtuelle à la prise en charge des troubles cognitifs et comportementaux*. Thèse de doctorat, Ecole Nationale Supérieure des Télécommunications, Février 2006. **Résumé** : Nous voyons dans cette thèse, un des emplois porteurs de la réalité virtuelle où des conférences sur ce thème ont déjà été initiées. L'emploi de la réalité virtuelle comme «calculateur» de certains troubles y sont décrites. Ce manuscrit présente aussi certains aspects de l'immersion utilisés pour traiter des troubles comportementaux.
- [49] Benjamin KUIPERS : A hierarchy of qualitative representations for space. *Lecture Notes in Computer Science*, 1404:337–350, 1998.
- [50] Harold Dwight LASSWELL : *Structure et fonction de la communication dans la société*. Encyclopaedia of the Sciences Social, 1948. **Résumé** : H.D. Lasswell présente sa décomposition de la communication qui a été utilisé pour la première fois dans le domaine de la communication de masse. Ce modèle populaire a ensuite été utilisé dans divers domaines. Ce modèle est connu par sa décomposition simple des problèmes de la communication.
- [51] Judith LAZAR : *La science de la communication*. Que Sais-je ?, Paris, France, 1992. **Résumé** : Explications très complètes sur l'ensemble de la communication. Ce livre assez court est une base de travail solide pour traiter le domaine de la communication.
- [52] Pascal LE MER : *Modèle de communication Homme-clone-Homme pour les environnements virtuels collaboratifs non-immerssifs*. Thèse de doctorat, Université des Sciences et Technologies de Lille, 2001. **Résumé** : Cette thèse présente un modèle d'architecture informatique, s'appuyant sur une typologie des gestes et sur le concept d'«interface de commande implicite par détection d'intention». L'objectif est de prendre en compte plusieurs moyens interactionnels et notamment de résoudre le problème de l'action et de la communication, inhérent aux EVC non-immerssifs. On y trouve plusieurs développements techniques comme la capture de mouvements ou l'animation faciale qui ont été intégré à la plateforme Spin-3D.
- [53] Thomas LEBARBÉ : Vers une plate-forme multi-agents pour l'exploration et le traitement linguistiques. *In TALN 2001*, GREYC - Université de Caen, 14032 Caen Cedex, France, July 2001. **Résumé** : Dans cet article, on retrouve une proposition de plate-forme multi-agents

pour l'expérimentation et le traitement linguistique. On y retrouve la description d'un agent APA au sein de cette plateforme avec une description des implémentations envisagées.

- [54] Eric LECOLINET : A brick construction game model for creating graphical user interfaces : The ubit toolkit proc. *In INTERACT'99.*, 1999. <http://citeseer.ist.psu.edu/300302.html>, **Résumé** : Présentation du toolkit Ubit qui permet de créer une interface. C'est un manuel explicatif sur les différentes méthodes visuelles.
- [55] Pascal LEMER, L. SOLER, D. PAVY, A. BERNARD, J. MOREAU, D. MUTTER et J. MARESCAUX : Argonaute 3d : A real-time cooperative medical planning software on dsl network. *In MMVR12, 12th Annual Medicine Meets Virtual Reality Conference*, Newport Beach, California, january 2004. Cette application du domaine médicale permet de faire intervenir différentes personnes géographiquement distantes autour d'une même table de travail. Les personnes impliquées partagent un point de vue et peuvent ainsi plus facilement prendre des décisions sur les sujets concernés.
- [56] Charles LENAY et Francois-David SEBBAH : La constitution de la perception spatiale. Approches phénoménologiques et expérimentales. *In Association pour la recherche COGNITIVE*, éditeur : *Intellectica (Intellectica)*, ISSN 0769-4113,, numéro 32 de 1, pages 45–85, 2001. INIST-CNRS, Cote INIST : 22461, 35400010029735.0020, <http://www.utc.fr/gsp/publi/Lenay01-LaConstitution.pdf>, **Résumé** : Cet article tente d'initier un rapport d'enseignement réciproque entre une étude scientifique de la perception, et la phénoménologie entendue comme une méthode philosophique, à propos de la constitution de la spatialité, et plus particulièrement de la profondeur spatiale. Cet article présente une étude expérimentale originale fondée sur des dispositifs techniques de suppléance perceptive.
- [57] David LERAY et Jean-Paul SANSONNET : Daft swing : une famille de composants dialogiques pour un agent conversationnel assistant. Mémoire de D.E.A., LIMSI/CNRS, Orsay XI, 2005. **Résumé** : Ce rapport présente une architecture permettant l'élaboration d'un agent conversationnel au sein d'une application. Il se veut capable de raisonner sur la structure et le fonctionnement de l'application et ainsi offrir un contexte d'explications pertinents.
- [58] Xin LI, Chieh-Chih CHANG, et S. K. CHANG : Face alive icons. *Journal of Visual Language and Computing*, Feb 2007.
- [59] J. LOMBARDI, H. STEARNS, J. GARGUS et M. SCHUETZE : The brie user experience : End-user distributed authoring via construction and deconstruction. *In the Fourth Conference on Creating, Connecting, and Collaborating through Computing*, 2006.

- [60] Marilyn M. LOMBARDI et Julian LOMBARDI : Croquet learning environments : Extending the value of campus life into the online experience. *In Croquet3D*. The university of Wisconsin-Madison, Division of Information Technology, 2005.
- [61] Matthias LUDWIG, Alexandre LENOIR et Pierre NUGUES : A conversational agent to navigate into mri brain images. *In Proceedings Interfaces 97, 6th International Conference*, pages 207–210, 28–30, Montpellier, Mai 1997. <http://cite-seer.ist.psu.edu/ludwig97conversational.html>, **Résumé** : L'agent conversationnel de P. Nugues a été employé dans l'exploration des images du cerveau. Les différentes zones du cerveau sont nommées. L'agent à l'aide de ces nominations peut comprendre et naviguer dans les reconstructions virtuelles.
- [62] J.F. MARCOTTE : Interactions en réseaux et communautés virtuelles, 2003. *La Communication Médiatisée par Ordinateur : un carrefour de problématiques*.
- [63] Jean-Claude MARTIN : *Le Guide de la communication*. Editions Marabout, Collection Vie Professionnelle, 1999. ISBN :2501032292, hachette :4028833, EAN :97825010322.
- [64] Jean-Claude MARTIN : *Coopération entre modalités et liage par synchrone dans les interfaces multimodales*. Thèse de doctorat, GET ENST Paris, 1994.
- [65] M. MAYBURY : Intelligent multimedia interfaces. *In Computational Linguistics*, pages vi + 405. MITRE Corporation Menlo Park, CA : AAAI Press and Cambridge, MA : The MIT Press, 1993. Paperbound, ISBN 0-262-63150-4, <http://cite-seer.ist.psu.edu/wittenburg93review.html>, **Résumé** : Ce papier montre les choix effectués en matière de système multimodal en entrée et en sortie. On y retrouve certaines de nos explications pour les améliorations. Mais la fusion est traitée de manière détaillée, et on y trouve les explications sur son implémentation dans le système COMET avec un système de désignation de type cybergloves.
- [66] Denis MCQUAIL : *Research on new communication technologies : barren terrain or promising arena ? shaping the future of communications*, pages 431–445. G.K. Hall and Co. Boston, MA, USA, 1987. ISBN :0-8161-1851-5.
- [67] M. MILGRAM : *Reconnaissance des formes, méthodes numériques et connexionnistes*. Collection 2ai, 1993. Armand Colin. **Résumé** : Livre généraliste sur la reconnaissance des formes. Il présente différentes approches et méthodes pour reconnaître des formes géométriques. L'auteur détaille notamment un exemple de reconnaissance par des réseaux de neurones.

- [68] Jean-Michel MOREL, Pablo MUSÉ et Frédéric SUR : Recherche dans les grandes bases de formes. *In Centre de Mathématiques et de Leurs Applications École Normale Supérieure de Cachan*, 61, avenue du Président Wilson 94235 Cachan Cedex, 2002. **Résumé** : Cet article présente une méthode d'appariement sur des formes contenues dans les images. L'approche retenue permet une décision sans paramètre. Des résultats expérimentaux sur la recherche de logos sont présentés.
- [69] Alex MUCCHIELLI : *La nouvelle communication : Épistémologie des sciences de l'information-communication*. Armand Colin Edition, 2000. ISBN :2200016964. **Résumé** : Approches théoriques sur les nouvelles technologies de la communication et de l'information. Livre généraliste qui fournit des explications sur l'évolution des différents modèles de communication.
- [70] Theodore M. NEWCOMB : An approach to the study of communicative acts. *In Psychological Review*, 1953. 60, Heft 6, S.393–404, **Résumé** : Ce livre présente un rassemblement de toutes ses recherches précédentes, il décompose surtout les problèmes de communication permettant une compréhension aisée de son modèle.
- [71] Laurence NIGAY et Joëlle COUTAZ : Espaces conceptuels pour l'interaction multimédia et multimodale. *In Spécial Multimédia et Collecticiel, AFCET Hermes Publ., Vol 15(9)*., pages 1195–1225, 1996. **Résumé** : Les avancées rapides de la technologie se manifestent par l'apparition de nouveaux termes parfois sans rigueur. Cet article fait le point sur la terminologie et sur les référentiels applicables à la conception d'interfaces multimédia et multimodales.
- [72] Laurence NIGAY et Frédéric VERNIER : Représentations multiples d'une grande quantité d'information. *In Proceedings IHM 97*, Laboratoire CLIPS-MAG, Equipe IIHM, Grenoble, France, 1997. **Résumé** : Cet article énonce certains principes relatifs à la conception et à la mise en œuvre de techniques de visualisation et plus généralement d'interaction dans le cas d'une grande quantité d'informations. L'article propose l'application d'une méthode de conception descendante fondée sur une étude de l'utilisateur et de son activité. Il traite plus particulièrement de la représentation multiple et présente une technique de visualisation à plusieurs facettes.
- [73] P. NUGUES, C. GODÉREAUX, P.O. EL GUEDJ et F. REVOLTA : A conversational agent to navigate in virtual worlds. *In Susann LUPERFOY, Anton NIJHOLT et Gert Veldhuijzen van ZANTEN, éditeurs : Proceedings of the eleventh Twente Workshop on Language Technology, TWLT 11, Dialogue Management in Natural Language Systems*, pages 23–33, Enschede, 1996. Universiteit Twente. <http://cite-seer.ist.psu.edu/nugues96conversational.html>, **Résumé** : On y présente le prototype de l'agent conversationnel de P.Nugues. Cet agent permet

- une navigation assistée. Sa méthode d'interaction par la voix permet une meilleure immersion de l'utilisateur dans le monde virtuel. les auteurs présentent l'architecture de l'agent et de ses différents composants : l'analyseur linguistique, un raisonneur géométrique, la reconnaissance vocale et le synthétiseur de voix.
- [74] Dimitris PAPADIAS et Max J. EGENHOFER : Algorithms for hierarchical spatial reasoning. *GeoInformatica*, 1(3):251–273, 1997.
- [75] Charles J. PETRIE : Agent-based engineering : the web and intelligence. *In IEEE Expert*, Stanford Center for Design Research, Décembre 1996. <http://www-cdr.stanford.edu/NextLink/Expert.html>Petrie96, **Résumé** : Cet article présente un agent compatible avec internet qui utilise un «KQML-like» pour récupérer et traiter les informations. On y décrit les problèmes survenus avec les notions d'agents et les protocoles d'internet. Il fournit une explication sur le concept d'autonomie d'un agent.
- [76] T. PFEIFFER et M.E. LATOSCHIK : Resolving object references in multimodal dialogues for immersive virtual environments. *In Virtual Reality, 2004. Proceedings. IEEE*, page 35, Bielefeld University, Germany, March 2004. ISSN : 1087-8270, ISBN : 0-7803-8415-6, **Résumé** : Ce papier décrit les grandes lignes d'un système de résolution de références en multimodalité. Le système est basé sur des concepts humains. On y retrouve une étude sur l'implémentation de certains concepts cognitifs au niveau du système virtuel.
- [77] Virtual REALITY et Scientific Visualization in ARCHAEOLOGICAL RESEARCH : Collaborative visualization of an archaeological excavation. *In Virtual Archaeology between Scientific Research and Territorial Marketing*, Arezzo, Italy, Novembre 2000. VAST 2000. **Résumé** : C'est un article sur le système ARCHAVE constitué d'une CAVE. On y montre les apports de la réalité virtuelle dans la reconstruction d'un site archéologique. Il tente de fournir un support pour la compréhension de la fonction des objets en fonction de son emplacement dans les fouilles archéologique.
- [78] Jochen RENZ et Bernhard NEBEL : Spatial reasoning with topological information. *Lecture Notes in Computer Science*, 1404:351–365, 1998.
- [79] Nadine RICHARD : *Description de comportements d'agents autonomes évoluant dans des mondes virtuels*. Thèse de doctorat, ENST, 46, rue Barrault - 75013 Paris, France, 2001. **Résumé** : Ces travaux présente une plate-forme multi-agents qui implique que les messages échangés et toutes les communications possibles se produisent exclusivement entre les agents. Les informations ainsi transmises concernent des états simples ou des paramètres plus ou moins complexes. Nous ne pouvons parler ici des différentes composantes de la communication comme pour les autres exemples étudiés car nous sommes dans un contexte différent.

- [80] Nadine RICHARD, Philippe CODOGNET et Alain GRUMBACH : The inviwo toolkit : Describing autonomous virtual agents and avatars. *In Actes du workshop Object-oriented VRML de la conférence VRML'98*, 46 rue Barrault 75013 Paris, France, 8 rue du Capitaine Scott, 75015 Paris, France, Septembre 2001. ENST, LIP6.
- [81] Richard RIEDL, Thesera BARRETT, Janice ROWE, William VINSON et Susan WALKER : Sequence independent structure in distance learning. *In Conference on Computers and Learning*, Department of Leadership and Educational Studies, Reich College of Education, Appalachian State University, Boone, NC 28608, USA, 2000. <http://www.activeworlds.com/edu>.
- [82] Matilda White RILEY et John White RILEY : A sociological approach to communications research. *Public Opinion Quartely*, 15(3):445–460, 1959.
- [83] I. ROJDESTENSKI, D. MODJESKA, F. PETTERSSON, M. ROJDESTVENSKAIA et P. GUSTAFSSON : Sequenceworld : a genetics database in virtual reality. *In Information Visualization, Proceedings. IEEE International Conference*, pages 513–517, London, Angleterre, September 2000. Departement of Plant Physiol., Umea Univ. INSPEC Accession Number : 6663866, Digital Object Identifier : 10.1109/IV.2000.859805, ISBN : 0-7695-0743-3.
- [84] Nicolas ROUGNON et Henri MAÎTRE : Eléments pour la reconnaissance de formes tridimensionnelles déformables. application à l'imagerie biomédicale. *In Travaux universitaires – ENST-Paris, FRANCE*, 1993. INIST-CNRS, Cote INIST : T 88657, Résumé : Article sur la reconnaissance de formes par déformations. Ils se sont placés dans un cadre déterministe continu et adoptés une approche régularisante non-paramétrique de type snake. Un modèle multidimensionnel déformable, dénommé stabilisateur membrane/plaque-mince sous pression ou g-snake, est élaboré et caractérisé analytiquement.
- [85] Jean SALLANTIN : *Les agents intelligents : Essai sur la rationalité des calculs*, volume 1. Hermes Sciences Publications, Octobre 1998. ISBN : 2866016432.
- [86] Marilyn C. SALZMAN, C. DEDE et R.B. LOFIN : Learner-centered design of sensorily immersive microworlds using a virtual reality interface. *In Human Factors and Ergonomics Society Annual Meeting*, volume 39, pages 486–490, San Diego, CA, United States, Octobre 1995. Résumé : Newton World est l'un des premiers mondes virtuels créé dans le but de comprendre l'impact de la gravité dans le monde réel. Il fournit un outil de visualisation des effets de la gravité sur les objets. Ces derniers ont différentes masses et la manipulation rapide des différentes variables du monde nous fournit une appréciation du rapport masse/gravité.

- [87] Henry SAMIER : *La recherche intelligente sur internet et sur l'intranet*. Hermes Sciences Publications, Septembre 1999. 2^{ème} édition, ISBN : 2746200708. **Résumé** : Présente la mise en place et le fonctionnement du programme IntelliShoper. Il décrit l'agent qui permet d'aller chercher des informations sur Internet suivant les demandes de l'utilisateur.
- [88] G. SANTORO : What is computer mediated communication? *In* Z. BERGE et M. COLLINS, éditeurs : *Computer mediated communication and the On-line Classroom in Distance Education*, volume 1, New Jersey, 1995. Hampton Press. Overview and Perspectives.
- [89] Ralph SCHROEDER, Avon HUXOR et Andy SMITH : Activeworlds : geography and social interaction in virtual reality. *In Futures*, Département of technology and Society, University, S-412 96 Gothenburg, Sweden, 2001. Elsevier Science Ltd. **Résumé** : Active World (ou Alpha World) est l'un des projets de monde collaboratif virtuel les plus anciens. Ce monde s'est rapidement développé pour atteindre aujourd'hui une taille comprenant près de 60 millions d'objets virtuels visualisables. Il est devenu un monde virtuel immense où les créations sont nombreuses.
- [90] Lucien SPEZ : *La communication*. Que Sais-je ?, Paris, France, 1991. ISBN : 2130542131. **Résumé** : Livre fournissant des informations et une approche différente de Judith Lazar sur le domaine de la communication. Il apporte des informations annexes assez pertinentes sur la communication.
- [91] Claude SHANNON et Waren WEAVER : Théorie mathématiques de la communication. Bibliotheque du CELP, 1945. **Résumé** : Modèle considéré comme une référence dans de nombreux domaines traitant du signal.
- [92] R. SHARMA, M. YEASIN, N. KRAHNSTOEVER, I. RAUSCHERT, G. CAI, I. BREWER, A. MACÉACHREN et K. SENGUPTA : Speech-gesture driven multimodal interfaces for crisis management, 2003. <http://cite-seer.ist.psu.edu/572016.html>, **Résumé** : Les interfaces sont parfois difficile à employer et présente souvent un temps d'apprentissage. Cet article présente Crisis management, un système multimodal qui traite les gestes et la voix. Le contexte de ce système est particulier. Il se centre sur la minimisation des risques. Il souhaite enregistrer toutes les étapes d'un projet pour pouvoir les suivre par la suite depuis le commencement par une visualisation notamment par une restitution de toutes les informations du projet notamment les actes de communication entre les acteurs du projet.
- [93] M. SILVA, S. ZHANG, C. DEMIRALP et D. LAIDLAW : Visualizing diffusion tensor volume differences. *In IEEE Visualization 01 Proceedings*, Brown University Computer Science, Providence, RI, USA,

1999. <http://citeseer.ist.psu.edu/silva01visualizing.html>, **Résumé** : Cet article présente un des intérêts grandissant du domaine médical pour la réalité virtuelle. Il permet une visualisation plus simple et élargie des informations de tensions au niveau du cortex cérébral.
- [94] David A. SMITH, Alan KAY, Andreas RAAB et David P. REED : Croquet 3d : A collaboration system architecture. In *Croquet3D*, 104 So Tamilynn Cr., Cary NC, 27513, 2003. **Résumé** : Open Croquet est une plate-forme orientée vers la création d'applications multi-utilisateurs collaboratives en ligne. Une grande partie des recherches sur cet environnement s'est portée sur l'architecture permettant de supporter une communication avec différents supports. Il présente des implémentations orientées vers le partage de ressources entre les utilisateurs.
- [95] Timo SOWA, S. KOPP et M. LATOSCHIK : A communicative mediator in a virtual environment : Processing of multimodal input and output. In *Proceedings of the International Workshop on Information Presentation and Natural Multimodal Dialogue*, 345, pages 71–74, Verona, Italy, 2001. <http://citeseer.csail.mit.edu/sowa01communicative.html>, **Résumé** : Cet article présente un système multimodal de communication avec un agent. Les travaux se sont portés sur les gestes et son emploi synergique voire redondante avec la voix. Une partie de cet article présente la production de gestes et paroles coordonnés.
- [96] Timo SOWA et Ipke. WACHSMUTH : Coverbal iconic gestures for object descriptions in virtual environments : An empirical study, 1999. Technical Report, <http://citeseer.csail.mit.edu/sowa99coverbal.html>.
- [97] Erik STROMMEN : Children's use of mouse-based interfaces to control virtual travel. In *CHI '94 : Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 405–410, Boston, Massachusetts, United States, 1994. ACM Press. <http://portal.acm.org>, ISBN : 0-89791-650-6, **Résumé** : Cet article montre les performances des enfants utilisant différentes interfaces. Il met en évidence certaines actions intuitives. Une des conclusions sur la souris indique qu'elle est d'un emploi assez intuitif.
- [98] Louise TREMBLAY : *La qualité de la langue et les médias écrits*, pages 13–19. UQAM-Revue québécoise de linguistique théorique et appliquée – Le Français moderne, 2000. **Résumé** : Article intéressant sur la longueur des phrases, avec des explications sur les règles de Flesh et Gunning avec une conclusion sur le rapport entre la longueur des phrases et l'impact. Les médias écrits constituent un terrain d'observation idéal pour des linguistes. L'auteur analyse la qualité de la langue, sa définition et sa description, La qualité de la langue est en effet une notion fluide, difficile à saisir. En tant que linguiste, il énonce des problèmes sur l'accès à la qualité sans faire preuve de « rectitude linguistique » ou, à l'inverse, de laxisme à outrance.

- [99] Lionel ULMER : Vreng : Analyse detaillee - module interface. Rapport technique, ENST, 1998.
- [100] Johann VANDROMME, Samuel DEGRANDE, Patricia PLÉNACOSTE, Luciana PROVENZANO et Christophe CHAILLOU : Un système de visite virtuelle pour un bureau 3d. *In Interaction Homme-Machine*, Montréal, Canada, April 2006. **Résumé** : Cet article propose un système de navigation dans un objet 3D. Il est inspiré de la notion de «world in miniature» et utilise un système de boule de cristal pour afficher l'environnement de l'utilisation et ainsi aider à en comprendre la topologie.
- [101] Eric VARADARADJOU : Les communications dans les mondes virtuels. Rapport technique, ENST-Paris, 2003. **Résumé** : Chaque monde virtuel possède ses possibilités de communication, nous analysons ici les différentes communications existantes dans divers mondes virtuels.
- [102] Eric VARADARADJOU : Les langages de description dans les mondes virtuels. Rapport technique, ENST-Paris, 2004. **Résumé** : Ce rapport etudie les possibilités de différents langages de description de scène dans le but de trouver une solution de structure de données pour VREng.
- [103] Eric VARADARADJOU, Philippe DAX et Alain GRUMBACH : Communication améliorée dans les mondes virtuels. *In Conférence : Association Française de Réalité Virtuelle AFRV'06*, pages 83–89, INRIA Rocquencourt, France, November 2006.
- [104] Eric VARADARADJOU, Philippe DAX et Alain GRUMBACH : Improved communication in virtual world. *In Conférence : Association Française de Réalité Virtuelle, IEEE France, VRIC'06*, pages 103–111, Laval, France, April 2006.
- [105] Richard C. WATERS, David B. ANDERSON, John W. BARRUS, David C. BROGAN, Stephan G. McKeown MICHAEL A. CASEY, Tohei NITTA, Ilene B. STERNS et William S. YERAZUNIS : *Diamond Park and Spline : A Social Virtual Reality System with 3D Animation, Spoken Interaction, and Runtime Modifiability*, volume 6, 4, pages 461–480. Presence, Août 1997.
- [106] Kenneth WAUCHOPE : Multimodal interaction with a map-based simulation system, 1996. <http://cite-seer.ist.psu.edu/wauchope96multimodal.html>, **Résumé** : Ce rapport décrit «InterLACE», un système multimodal combinant l'interface militaire d'une simulation de plan et le langage naturel. Il contient surtout une étude sur la communication humain-humain dans un contexte virtuel. Le système analyse tous les informations du monde virtuel. Par sa connaissance complète du monde et de l'interface, le système répond rapidement aux requêtes de l'utilisateur.

-
- [107] SuperScape 3D WEBMASTER : *User's guide*. Virtual Reality Software, 1997. <http://www.thirddimension.co.uk/pages/products1.htm>,
Résumé : Ce guide comporte le tutoriel permettant l'apprentissage progressif de 3D Webmaster. Il fournit les différentes utilisations et outils disponibles. Le langage SCL y est décrit de manière précise avec de nombreux exemples.
- [108] B. H. WESTLEY et M. S. MACLEAN : A conceptual model for communications research. *Journalism Quarterly*, 34(4):31–38, 1957.
- [109] Mason WOO, Jackie NEIDER, Tom DAVIS et Dave SHREINER : *Guide Officiel OpenGL 1.2*. Campus Press Référence, Juin 2000. ISBN : 2744008419. Résumé : Les différentes subtilités de développements faites lors de notre implémentation n'aurait pu se faire sans une première approche à travers ce livre complet d'OpenGL.
- [110] Guy WRIGHT : Le guide des écrans. Rapport technique, THG Web, 2001. Résumé : Etude générale sur les écrans cathodiques et LCD, études complexes et très complète pour comprendre les possibilités des écrans.
- [111] Jean Caelen YANNICK FOUQUET : Deux plates-formes pour l'expérimentation d'un agent conversationnel anime. *In Laboratoire CLIPS-IMAG*, 2005. Résumé : L'article présente deux plates-formes d'expérimentation de dialogue homme-machine, avec pour exemple un agent conversationnel expressif et gestuel dans le domaine du commerce électronique.
- [112] Nick YEE, Jeremy N. BAIENSON, Mark URBANEK, Francis CHANG et Dan MERGET : The unbearable likeness of being digital : The persistence of nonverbal social norms in online virtual environments. *In CyberPsychology and Behavior*, Department of Communication, Stanford University, 2007. Résumé : Les mondes virtuels sont nombreux, Ces environnements ont été une plateforme d'étude pour analyser certains comportements. L'étude porte sur Second Life où les auteurs analysent la distance entre les avatars suivant différents facteurs. A l'aide des résultats obtenus et des comparaisons entre des résultats dans le réel et le virtuel les amènent à des discussions sur les implications possibles.

Index

- 3D Webmaster, 53
- Active World, 62, 74, 78
- Agent, 96, 99, 107, 123
 - Agent de recherche, 97
 - Agent intelligents, 98
- Améliorations déictiques, 112, 138
- Améliorations textuelles, 111, 128
- Améliorations visuelles, 112, 143
- Chaffee et Berger, 22
- Chaillou Christophe, 47, 65, 68, 113, 115
- Communication
 - Définitions, 22
 - Evolution technologique, 23
- Communication interpersonnelle, 25
- Croquet, 64, 78
- Diamond Park, 59, 116
- Dimension de contrôle, 33
- Dimension perceptive, 33
- Dive, 50, 65, 78
- Gerbner, 32, 36, 71, 81, 82, 86, 107
- Habbo, 66, 78
- Langage SCL, 53, 54
- Lasswell, 30, 50, 78, 80, 85
- Le Village 3D, 73
- LEA-LIMSI Embodied Agent, 70
- Limitations des MV, 47, 91
- Méthodes d'ajout, 110, 128
- Méthodes de prévention, 113, 147
- Méthodes de réponse, 114, 150
- Martin Jean-Claude, 70, 100, 167
- McQuail, 22, 24
- Modalité, 100, 102, 103
- Mode picking, 135, 140, 172
- Multimodale, 91, 104, 107
 - Fission multimodale, 107
 - Fusion multimodale, 101
- Newcomb, 34, 82, 94, 155
- Newton World, 47
- Nigay Laurence, 100, 141
- Niveaux de communication, 22
- Nugues Pierre, 55, 105
- Off-screen rendering, 135, 144, 177
- Polysémie, 90
- Position automatique, 141, 175
- Réalité virtuelle
 - Dispositifs immersifs, 42
 - Dispositifs non-immersifs, 43
 - Dispositifs semi-immersifs, 43
 - Les communicants, 49
- Riley et Riley, 37, 81, 93
- Second Life, 72, 78
- Sequence World, 45
- Shannon et Weaver, 27, 67, 80, 86
- Spin-3D, 68
- Ulysse, 55, 76
- VR4all, 78
- VREng, 58, 118
- Westley et Mac Lean, 35
- XML, 97, 120, 122, 129